A DESIGN OF A PERSONNEL TRAINING INFORMATION SYSTEM FOR THE U. S. MARINE CORPS

Robert Anthony VanHouten

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

A DESIGN OF A PERSONNEL
TRAINING INFORMATION SYSTEM
FOR THE U.S. MARINE CORPS

by

Robert Anthony VanHouten, Jr.
Robert Paul Hansen
Robert Eugene Sonnenberg, Jr.

June 1975

Thesis Advisor: N. F. Schneidewind

Approved for public release; distribution unlimited.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Design of a Personnel Training Information System for the U. S. Marine Corps		5. TYPE OF REPORT & PERIOD COVERED Master's thesis; June 1975
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Robert Anthony VanHouten, Jr. Robert Paul Hansen Robert Eugene Sonnenberg, Jr.		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE June 1975
		13. NUMBER OF PAGES 161
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

training information system
mini-computer training information system
Marine Corps training information system

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A personnel training information system is designed for use on minicomputers within the operational forces of the U. S. Marine Corps. The problem of maintaining and accessing training information is analyzed. A computerized system and the characteristics it must possess are defined. The data base is formulated, input requirements are stated, and software support is defined. Alternatives for software support, including three existing hardware and software minicomputer systems, are

20. (cont.)

discussed. A concept for a minicomputer network is developed. The specific requirements for the exchange of training information within the network are analyzed. Emphasis is placed on integration with existing hardware and software systems.

A Design of a Personnel
Training information System
for the U.S. Marine Corps

by

Robert Anthony VanHouten, Jr.
Major, United States Marine Corps
B.S., Georgia Institute of Technology, 1965

Robert Paul Hansen
Captain, United States Marine Corps
B.A., University of Wisconsin, 1968

Robert Eugene Sonnenberg, Jr.
Captain, United States Marine Corps
B.S., United States Naval Academy, 1970

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

from the

NAVAL POSTGRADUATE SCHOOL
June 1975

ABSTRACT

A personnel training information system is designed for use on minicomputers within the operational forces of the U. S. Marine Corps. The problem of maintaining and accessing training information is analyzed. A computerized system and the characteristics it must possess are defined. The data base is formulated, input requirements are stated, and software support is defined. Alternatives for software support, including three existing hardware and software minicomputer systems, are discussed. A concept for a minicomputer network is developed. The specific requirements for the exchange of training information within the network are analyzed. Emphasis is placed on integration with existing hardware and software systems.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Most of the software systems used for data processing within the Marine Corps are resident on the IBM S/360 series of computers. They are designed to aid the high level manager in his decision making. Input to these systems is originated at the lower levels. The lower level commander does not, however, possess the capability of tapping the vast amount of data that is available to satisfy his information requirements and therefore, must resort to manual systems.

The Marine Corps is also faced with the problem of providing data processing support when a unit is deployed for an extended period of time. The present hardware configuration is limited in its mobility. Data processing support should be available to the local operational units when deployed because reverting to a manual processing is becoming increasingly more cumbersome.

State-of-the-art hardware and software systems exist that, though costly, will provide the needed support. This hardware and software exists in the form of minicomputers and data base systems. This thesis, then, will develop requirements for a minicomputer system to solve the problem of providing a deployable information retrieval capability for the local organizations.

First, an example of a local information system must be developed. Section II discusses the problem of maintaining and accessing training information. Emphasis is placed on integrating this function with existing personnel management systems.

Section III defines the envisioned computerized system

and the characteristics it must possess. The data base is formulated, input requirements are stated, and software support is defined.

Section IV addresses further software support required and alternatives that exist for obtaining this support. A general discussion of three existing hardware and software systems is presented.

Sections V and VI discuss the exchange of information through the use of a minicomputer network. Section V is a general discussion of networking considerations and implications. Section VI transforms these general ideas into specific applications based on the training information system developed in Sections II and III.

Finally, Section VII presents conclusions and recommendations, respectively, based on the research conducted to formulate the thesis.

II. SYSTEM REQUIREMENTS

A. BACKGROUND

The Marine Corps commander, like any other manager, requires an information base upon which to make decisions. In the past such information was collected via various manual reporting systems. Because of processing, quantity and transmission limitations inherently associated with such a manual system, the commander's information was often incomplete, outdated, and (in worst cases) incorrect.

In order to improve the commander's decision making capability, the Marine Corps has developed automated data processing and reporting systems. The current systems rely on extensive input at the lowest level of reporting units. The information derived becomes a vital part of the higher echelon organization management function. However, because of equipment and system design limitations, it is often not possible to provide relevent information to the lowest level reporting unit for incorporation into their management process. Besides the obvious disadvantage of requiring the base unit to maintain a dual information system, the present system often has another very detrimental effect. A sense of frustration and resistence to the unresponsive system is created resulting in an unmeasurable but significant reduction in overall mission effectiveness.

Headquarters Marine Corps contracted with Informatics Incorporated to undertake a study on the present system and propose possible improvements. The study was conducted during the period February, 1973 to September, 1974. Its official title was "Data Management Device Requirements (DMDR) Study." A summarization of the Study's objectives follows:

1. To identify data management and teleprocessing requirements which could be profitably automated at lower echelon reporting units.

2. To identify and size a hardware system for field test.

3. To determine the feasibility of processing equipment and recommend alternative devices for consideration by the Marine Corps.

4. To recommend an initial implementation plan for test of the recommended system.

As a result of their efforts, Informatics developed a set of conclusions and recommendations. Stated in a summarized form the conclusions are:

1. There are management problems within the Marine Corps which can be remedied through the use of automated data devices. The particular problems are deployment logistics management, individual training, decentralized Supported Activity Supply System (SASSY) entry and the Joint Uniform Military Pay System (JUMPS)/Marine Corps Manpower Management System (MMS).

2. There exist devices which are compatible with garrisoned and deployed modes of operation of Fleet Marine Force (FMF) units.

3. A basic data management device (DMD) can satisfy the requirements of many of the identified application areas. It consists of

   64K 16 bit word processor
   keyboard entry
   visual display
   printer
   disk storage
   cassette tape output

4. There exist some applications which are specialized

but which are of such magnitude that an automated system is justified. These include Combat Essential Equipment Status/Maintenance management, centralized SASSY entry and Marine Aviation requirements.

5. The acquisition of Data Management Devices (DMD) and implementation of the system will result in significant manpower savings and improvements in the operation of the FMF.

The recommendations of the study, again in summary form, are:

1. That several devices with the above stated characteristics be selected for test.

2. That the devices be tested in different types of FMF organizations.

3. That the devices be tested in a deployed situation in order to evaluate mobility, reliability and maintainability.

4. That a multi-functional application on one device be tested at a minimum of one location.

5. That cost benefit data be collected as a part of the various tests.

6. That test results be utilized to develop performance criteria for a final recommended device.

B. OBJECTIVES

This thesis is intended to develop the ideas and suggestions presented in the DMDR Study. The objective is to develop and present an automated system for the management of a typical training program at the battalion or squadron level. This process involves a detailed analysis of the present system. This includes a definition of requirements in the form of necessarry output, data base content, and inputs. Further, the thesis will define the

set of application programs required for the system. Based on this information the nature of the required system will be discussed to include specification of types of hardware and software.

Any proposed system in the personnel management area must be integrated with the existing personnel reporting system, the Manpower Management System (MMS). This is necessary in order to make use of existing data and avoid any unnecessary duplication of data elements. Also inherent to integration is the capability of building on a proven system. This proposed integration will be further discussed throughout the thesis.

C. TRAINING PROBLEM

Marine Corps Order (MCO) P1510.26, Unit Level Training Management, is a logical starting place for the definition and analysis of the Marine Corps training program. The order states the purpose of Marine Corps training is to "develop skilled forces-in-readiness prepared at all times to carry out any mission which may be assigned." Further, the order establishes that the training necessary to accomplish this purpose is the responsibility of the individual commander and is accomplished mainly through the use of resources organic to the unit. Presently, training management, mostly record keeping, is accomplished at the company level (with supervision from battalion) and at the squadron level. In this thesis local unit connotes the squadron/battalion level.

Continuing, the order specifies that the Marine Corps training program is divided into the following major subprograms:

1. Individual training.

2. Unit training.

3. Fleet Marine Force training.

4. Reserve Forces training.

Individual training and unit training are of primary concern at the small unit level. These are the areas upon which the commander will design his training program.

In order to develop a sound training program the commander must establish a set of realistic training objectives based on a thorough examination of his training situation. This involves an analysis of training requirements. Requirements are generated by three major sources: higher headquarters, the command's mission, and the career training needs of individual marines in his command.

Higher headquarters requirements are specified by Marine Corps directives and directives of intermediate commands. The applicable Marine Corps Orders are specified in MCO P1510.26 and listed in Appendix A with discussion. The bulk of the requirements levied by Headquarters apply to all marines and hence can be generalized for a Marine Corps-wide program.

Mission requirements are derived from a unit's table of organization and any technical or doctrinal publications which apply to the particular unit. Hence, training to fulfill mission requirements differs from unit to unit and cannot be generalized.

Individual career training requirements are outlined in MCO 1510.2H, The Individual Training of Enlisted Marines. In order to fulfill these requirements, the Commander must insure that marines receive the training necessary for increased rank and responsibility. This effort involves

leadership and military occupational speciality (MOS) training. This training applies to all marines and can be generalized to a Marine Corps-wide program.

A discussion of the commander's analysis of training requirements is also pertinent to analysis of the training program. As mentioned previously, both requirements originating from higher headquarters and requirements concerning individual career training are of a gèneral nature. Hence, they constitute the basis for the general automated training system which is desired. As a starting point the same list of directives outlined above is used. Each order has been analyzed with the purpose of developing the necessary output required to accomplish the training function. Once output has been determined, the data elements required to generate the output are easily defined.

D. CATEGORIES OF TRAINING

The following paragraphs discuss types of training which have been considere The aim is to group specific training requirements into categories of training which can td for incorporation into a computerized training information system. hen be examined as candidates for computerization. The categories are as follows: by Marine Corps Orders.

1. General training, the amounts of which are established by Marine Corps Orders.

2. Local training, the types and amounts of which vary between commands.

3. Mission oriented training, the types and amounts of which depend on unit mission and individual Military Occupational Specialty (MOS).

4. Educational training providing additional professional and academic skills through participation in voluntary programs.

General training is particularly well adapted to computerization. In this area there is uniformity throughout the Marine Corps. The data is maintained and accessed similarly in every unit making it easy to anticipate storage and retrieval requirements in designing a training information system.

The record keeping tasks for local training programs are identical to those of the Marine Corps-wide programs. However, there exists a problem of variation in the data maintained by units. This implies that training records providing for all training must be either non-standard or of sufficient size to incorporate the requirements of all commands simultaneously. The former approach is more flexible and for a computerized system more efficient in its use of storage. Thus this local training requirement suggests a training record which can be defined at each unit.

Mission oriented training is less easily computerized. Classroom lectures, demonstrations, and on the job training are often conducted on an ad hoc basis. Individual attendance records are often not kept. Subjects taught often reflect current needs rather than adherence to a preestablished curriculum. Any attempt to computerize such training could well discourage it by introducing unnecessary record keeping requirements. The exception occurs when a specific training syllabus exists or is set up locally. In this case there is a record keeping requirement. Again, the number of individuals using the syllabus may at any one time be so small that the use of a computer would not provide

assistance in managing the program. This area should not, however, be dismissed from further consideration. The implementation of mission oriented training in a training information system should be a user option, the efficacy of which can be expected to vary between units. Computerization of this type of training information poses the same problem as local training.

Educational training information is a potentially worthwhile addition to a training information system. The capability to organize a wide variety of information pertaining to individual skills can be realized through the use of a data base system which incorporates information from several sources into one educational training record. The record could include everything from service schools to correspondence courses to off duty civilian education.

E. REQUIREMENT FOR GROWTH

The foregoing discussion of training records attempts to extract from pertinent orders the essential data which must be kept in a training system. It constitutes little more than computerizing manual systems in existence today. It does not however address the subject of what additional training management could be performed once the basic training records have been entered into the computer. For example, it is not difficult to save the record of a previous year instead of purging it. The historical records could be used to evaluate remedial training programs. A variety of statistical compilations could be made to diagnose weaknesses in a unit's training program. It is not the purpose of this thesis to suggest additional desirable capabilities. It is appropriate however to note that given a computerized training management system, additional data storage and processing requirements would be rapidly

generated. Therefore, in designing a computerized training management information system to satisfy existing needs, liberal allowances for future growth should be made. It is not clear what information should be kept in a historical record. Because of the relatively high turnover rate of personnel, much of the data kept would not be related to any individual within the unit. Therefore, trends or histories on individuals would not be easily retrieved. Only statistical type data collection would be easily integrated into the system. Without knowledge of what might be required in a history record, the contents of a historical record will not be defined. However, an allowance for future incorporation of a history file should be made.

F. TRAINING INFORMATION REQUIREMENTS

Training information requests may result in one of the following outputs:

1. Generation of reports

2. Generation of training schedules

3. Summarization of training status

Reports are defined as written training summaries prepared for formal submission to higher headquarters. Since formal reports differ between units any attempt to identify a detailed list of reports required is futile. It is however possible to identify certain types of reports which can be expected in any unit. This can be done by describing the reporting process in terms of attributes and attribute values. For example, each marine possesses the attribute Physical Fitness Test (PFT). Prior to taking the test the attribute value is 'not taken'. After the test has

been taken the attribute acquires the value 'passed' or 'failed'. Training reports can be conveniently subdivided into the following general classes.

1. REPORTS OF PERIODIC TESSTING- These reports summarize the status of cyclic training programs. The form of the report is a list of the type training (attribute) and the numbers having specified values of the attribute. This information requirement suggests individual records of training.

2. REPORTS OF TRAINING CONDUCTED- These reports list the training subjects (attributes) most recently conducted with the numbers attending each (attribute value). Training subjects range from MOS training to personal affairs training. This information requirement suggests a group record of training.

3. REPORTS OF SPECIAL TRAINING ACTIVITY- These reports describe the status of training programs in which no large segments of the unit are required to participate. This area includes participation in Weight Control Programs or progress in Marine Corps Institute (MCI) correspondence course training. This area consists of several additional attributes peculiar to each individual, and suggests individual records for special training.

Scheduling of training can also be considered in terms of attributes. Scheduling consists of finding all individuals with a specific attribute value and then applying some criteria to reduce this number to the size appropriate for the training to be scheduled. The attribute value may be, for example, all marines not having fired a rifle this calendar year. The criteria to be applied in automatically selecting some subset of these for a rifle range detail is a more troublesome problem. In many cases some supervisor or even the individual himself is consulted prior to the assignment. If assignments could be made

automatically much time could be saved. Unfortunately, a completely automatic system would probably result in the scheduling of individuals who are not available for training at the time. So for scheduling applications a method for overiding the automatic selection for specified individuals should be available. A list of the individuals excluded but otherwise eligible to be scheduled should be output following the output of the schedule. Selection algorithms should provide for such selection options as proportionately by section, rank or a variety of other individual characteristics.

Summaries of training status can be generated by retrieving from a data base a list of individuals possessing a particular attribute value. Unlike reports which are anticipated retrieval requests, the generation of these lists requires retrieval based upon virtually every attribute value either singly or in combination.

1. Attributes and Attribute Values

Attributes are used to identify a class of information. Attribute values are logical or numerical forms that an attribute can assume. The following attributes are appropriate in a training information system:

1. Type of training

2. Rank

3. Work section

4. Relevant personnel information

An example of an attribute value is test score

(Pass, Fail or Incomplete). Examples of attributes and their values are as follows:

ATTRIBUTES	VALUE
1st Plt.-Rifle range	Not Fired
HUMAN RELATIONS	INCOMPLETE
PFT	FAILED

2. Listing Options

The information requested can take two primary forms, summary and individual lists. For summary information the following should be provided:

ATTRIBUTE	ATTRIBUTE VALUE	NUMBER	PERCENT

The data represent the number of marines which satisfy the the attribute value. The percent is the relationship between the number satisfying the attribute value and the number having the attribute. The following example illustrates this:

MCI course	delinquent	5	20 %

This indicates a 20 % delinquency rate for lesson submission of MCI courses for those enrolled.

For lists of individual marines the following information is appropriate for listing.

NAME	RANK	ATTRIBUTE	VALUE

The requirement to generate lists of names can take the form of listing names possesing a particular attribute

value (or combination of values) or listing the attribute values for all individuals possesing a certain attribute. There are many possibilities for ordering the printed list. The following options for listing seem appropriate for the majority of cases and as such will be considered the only system requirement for listing of information.

1. Alphabetical

2. Rank
   Alphabetical

3. Section
   Alphabetical

4. Section
   Rank
      Alphabetical

5. Score, high to low rank listing requires recording and checking the date of rank for correct order. This is useful in an administrative system, but it is not required for training.

3. Training Record Output

Upon transfer of a marine, his entire training record should be forwarded to his next command. This output request and the resulting formatted output can be considered as a separate special requirement not covered by the general output system discussed above.

4. <u>Frequency of Output</u>

Initially, the number of output requests is expected to be small. As the system becomes more widely used this number can be expected to increase significantly. In no event should output of training information present a heavy load on the system. The following table summarizes the expected output request load.

	TRANSACTIONS	
	DAILY	PEAK
NAME LIST GENERATION	5	10
UPDATE RECORDS	50	500

Peak loads can, for the most part, be avoided. The unavoidable peak loads occur when rosters of training completed for large numbers of individuals reach the training office at the same time.

5. <u>System Response Requirements</u>

Periodic training information lists are not normally required on short notice. The generated lists for schedules or reports may take up to 24 hours to produce. Special lists may, however, be required on short notice. System response times of 1 minute may become a requirement as the user begins to depend more on the system for information.

6. <u>Method of Output</u>

The training effort is conducted by marines without extensive specialized training. In order to avoid excessive training costs, it will be necessary for any automated system to be of a self contained nature. By self contained

is meant a system requiring little training for the end user. The means of communication with the system is via natural or English like non-procedural language. The user is required to learn only a limited number of vocabulary words and a method of inputing data base parameters in order to utilize the system. Such a language implies a non-batch processing mode. An on-line communication device such as CRT, terminal or telytpe is required.

G. ADDITIONAL INFORMATION REQUIREMENTS

The training record of an individual contains information pertinent primarily to the training management of that individual but not necessarily restricted to the training area. In fact, this training information is merely a subset of the overall personnel individual record. Some of the information kept on individuals is presently computerized while some is not. Data kept on individuals is found in the service record book (SRB) or officer qualification record (OQR), individual training record (ITR), and the JUMPS/MMS personnel record. Many of the data elements of these three records are recorded in at least two of the records and is thus redundant. Since the primary concern of this thesis is computerizing the training records as much as possible, with the added benefit of reducing redundancy, it is imperative to investigate what data elements are presently in a computerized data base and whether they can be used to form the individual training record.

A careful analysis of data elements and user requirements was conducted to formulate reports for the local level by operating forces of the Marine Corps. The resulting reports, shown in Appendix B, were derived from the MMS data base, and supposedly provide the local manager

with necessary relevant information. The capability to generate special reports utilizing other data elements of the basic 1200 byte MMS records exists with the Mark IV retrieval system. MARK IV was acquired by the Marine Corps to facilitate ad hoc retrieval requests on the existing S/360 systems.

Many of the data elements used in these reports may also be effectively used for training management. These elements can be updated only through the unit diary reporting system, which is source data entry porcedure for JUMPS/MMS.

Because this data already exists, it could be used to create a header for the individual training record. The header would be secure from erroneous update and greatly enhance the creation of the training record by alleviating the need to re-enter the data to form a new training file. Also inherent in this procedure is the required integration into existing systems. The actual design of the header will be discussed at a later time.

H. AUDIT TRAILS

The entire process of data input and validation should be kept as simple as possible so as not to induce the wrath of the user. Inherent to input is data preparation and some form of source document. It is assumed that some manual method currently exists for storing source documents for future validation purposes after records are updated. The time frame for keeping such documents will vary depending on the type of training and the number of visual audits of records conducted by an individual. The computerized system should not disrupt the current method other than to possibly simplify it.

Errors occurring during input of training data should be detected by the operator by comparing the raw input data to a system produced printed copy of all inputs. The system should detect errors whenever possible by checking that inputs fall within the correct range of values, i.e. numerical entries must be within a certain range and alpha entries must be defined for that type of training.

Periodically, a visual audit of the computer training records should be conducted. If errors are found, then the raw source data can be referenced manually in order to update the computer record.

I. DATA ELEMENTS

In order to specify the elements which will make up the recommended data base a table has been generated. The table contains the following information:

1. General Subject And Reference- This specifies the Marine Corps Order and its title which is the basis for a group of data elements.
2. Data Element- This specifies the field title for the particular element.

Columns 3 and 4 contain codes as follows:

A-Alphabetic

N-Numeric

A/N-Alphanumeric

3. Date- This specifies if it is necessary to maintain a date associated with a particular data element.
4. Score-This specifies if it is necessary to maintain a numerical score of training completed for a particular data element.
5. Bits-Size of field
6. Characters-size of field

7. Report- This specifies if it is necessary to submit a report based on the data elements associated with a particular reference. The following codes apply:

Q-Quarterly

S-semiannual

A-Annual

8. Schedule- This specifies if the data elements associated with a particular reference are used in generating training schedules and lists.

9. Transaction Rates-This specifies the transaction rate for the input scheduling of data elements. It is based on 250 work days per year, five day work week less national holidays. The dimension of the data is transactions per man per day.

The table was developed by analyzing the pertinent Marine Corps Orders. A discussion of each of the orders is in Appendix A.

SUBJECT/ REFERENCE	DATA ELEMENTS	FIELD DEFN DAT	SCR	SIZE BIT	CHR	REP- ORT	SCHE- DULE	TRAN RATE
INDIVIDUAL	Code conduct/		N	1			X	.012
TRAINING	UCMJ							
OF ENLISTED	History							
MARINES	customs							
MCO 1510.2H	courtesies		N	1			X	.012
	Close order							
	drill		N	1			X	.012
	Interior							
	guard		N	1			X	.012
	First aid/							
	field							
	sanitation		N	1			X	.012
	Uniform							
	clothing &							
	equipment		N	1			X	.012
	NBC Defense		N	1			X	.012
	Service rifle		N	1			X	.012
	Service							
	pistol		N	1			X	.012
	Individual							
	tactics		N	1			X	.012
	Leadership	N	N		6		X	.040
	Swimming							
	qualification		N		2		X	-
	Gas mask size				2			
MILITARY	Primary MOS		N		4			
OCCUPATIONAL	Secondary							
SPECIALTY	MOS		N		4			
(MOS)	Teriary							
MCO P1200.78	MOS		N		4			
	Billet							

	MOS		N	4			
TROOP INFORMATION PROGRAM MCO 1510.25A	Alcoholism/ alcohol abuse	N		4		X	.008
	Equal opportunity citizenship	N		4		X	.008
	Personal conduct/ character	N		4		X	.008
	UCMJ	N		4		X	.008
	Personal affairs	N		4		X	.008
MARKSMANSHIP TRAINING MCO 3574.2E	Rifle	N	N	7		X	.008
	Pistol	N	N	7		X	.008
	Shotgun	N		4		X	.008
TRAFFIC SAFETY PROGRAM MCO 5100.19A	Driver improvement course	N		4		X	.004
HUMAN RELATIONS PROGRAM	First year program	N		4	X	X	-
	Second year program	N		4	X	X	-
	Third/ subsequent program	N		4	X	X	-
	Current program	N		6	X	X	.008
	Unit discussion leader			1	X	X	-

PHYSICAL	Pullups	N	N	6		X	-
FITNESS AND	Situps		N	3		X	-
WEIGHT	Run		N	4		X	.024
CONTROL MCO 6100.3F	Unit endurance		N	4		X	.024
	Weight control		N	4	X	X	-
DRUG ABUSE	Initial instruction	N		4		X	-
CONTROL MCO 6710.1B	Overseas instruction			4		X	-
MARINE CORPS INSTITUTE MANUAL	Course titles(s)	N	N	10		X	-

TOTAL TRANSACTIONS PER MAN PER YEAR =.356.

Recommended Data Elements

Table I.

III. SYSTEM DESIGN

A. DATA BASE DESIGN

A few attempts to formulate and create a training information system have been made in the Marine Corps, according to the Marine Corps Automated Data Systems Plan Fiscal Years 1975-1980, none of which were standard and none of which were comprehensive. If a standardized training information system is to exist, a comprehensive, easily manipulated, integrated data base must be designed. Since training is itself only a part of the overall administrative function, the data base must be designed to facilitate the growth which will occur when other administrative functions are computerized. If the data base is developed with both the user and the processing in mind, it should prove to be acceptable for use in the Marine Corps.

In order to make effective use of both internal and auxiliary storage, a method of designing the data base to accomplish this task is proposed. The design considers the diversity of record types maintained on individual marines due to rank, duty, or special qualification and the possibility of future expansion of the data base to encompass many of the personnel administrative functions at the squadron/battalion level. The type of processing required by the application programs, mainly retrieval, will also be considered. For purposes of illustration, ancillary records will be developed to show their use and applicability.

The data base will consist of the following files:

1-Name/Rank File

2-Master File

3-Ancillary Aviator/NFO File

4-Ancillary Enlisted File

Each file will consist of fixed length records. Some of the files will be generated from existing Force Information Manpower Files (FISMNPWR) while some will be generated locally. Some of the benefits gained by using such a data base structure include more records per physical block, ease of update and reduced storage requirements. Some disadvantages include a more complicated data base management system and more complicated programs to generate lists and schedules, i.e. multiple file accesses to determine eligibility for training such as schedules for essential subjects training.

1. Description of Files

The following narrative describes the record layouts of each of the files in the data base. The field number, field name, field type, and field length in characters is given.

a. Name/Rank File

Each record in this file has the following description:

FIELD NO.	FIELD ID	FIELD TYPE	FIELD LENGTH
1	Last Name	A	15
2	Initials	A	2
3	Rank	A/N	2
4	Section/Company-platoon (code)	N	4 bits
5	Address of Next Alphabetical Record	N	11 bits
6	Print Indicator	N	1 bit
	Total Length of Record		21 ***

The name field will be used for double checking during update. The name, initials, and rank fields are obtained from the FISMNPWR record. The section/company-platoon code is used to aid the user in disseminating reports and locating individuals. Allowing 11 bits for a pointer permits up to 2048 records to be addressed. The print indicator bit is used for various processing tasks to be subsequently discussed. Records are stored by individual identification number (data base number), which is used to index the file.

b. Master File

This file contains data common to all marines. It consists of a header, derived from the FISMNPWR record, and a trailer containing training data, which is updated locally. The record description appears in Appendix C.

Fields 1-19 of the record are presently given to the squadron/battalion as the Command Personnel Report (Alpha) and the MOS and Location Report generated from the MMS data base. If the squadron/battalion has this information, they may run the reports as often as needed. Presently, the S/360 installations guarantee only one run a week and at most six copies of the reports. The S/360 installations also must distribute these reports. These limitations would be eliminated if the squadron/battalion had the capability to generate their reports.

No local updating of the header information (fields 1-25) should be allowed. This information is derived from the MMS data base and subject to the restrictions of the unit diary reporting system. Allowing local update of these fields would result in two similar files with different information for the same fields. Any

benefits from integration with the existing system would be lost.

The total length of the master record is 198 characters as is shown in Appendix C.

c. Aviator/NFO File

A good example of a ancillary file is an Aviator/Naval Flight Officer File containing information on a few individuals at the local level. Although this information could be included in the master record, a great deal of storage would be unused. This is because of the restriction of fixed length records. Enough space would have to be allocated in each record to contain this information. Since only a few marines will have any entries in these fields, the remaining marine's records would contain blank entries. Therefore, a ancillary file, the Aviator/NFO Ancillary File, will be used to save storage space.

This data presently exists on the FISMNPWR record. If the data were present at the local level, better planning and faster update could be accomplished. The format of the record is:

FIELD NO.	FIELD ID	FIELD TYPE	FIELD LENGTH
1	Aviation Update (yymmdd)	N	6
2	Total Pilot Hours	N	6
3	Pilot/NFO Hours Last Five Years	N	6
4	Total		
	Jet Hours	N	6
	Helicopter Hours	N	6
	Transport Hours	N	6
	Plane Commander Hours	N	6
	A/C Flown (Total of 5)	A/N	55
	Model (5)		
	Hours/Year (6)		
5	Date Designated Military Pilot	N	6
6	Back Pointer (DBN)	N	16 bits
	Total Length of Record		104 ***

d. Enlisted Ancillary File

Because certain areas of training are given only to enlisted personnel an ancillary record containing this information is proposed. Again, the motivation behind this approach is reduction of required storage. The enlisted ancilla ry record format is:

FIELD NO.	FIELD ID	FIELD TYPE	FIELD LENGTH	
1	Essential Subjects Training			
	Ten Subjects (Pass/Fail)	N	10	
2	Troop Information Program			
	Five Subjects-dates taken (mmyy)	N	20	
3	Leadership Training (NCO Only)			
	Date of Last Training (mmyy)	N	4	
	Stage of Training	N	2	
6	Back Pointer (DBN)	N	16	bits
	Total Length of Record		37	***

2. File Creation and Modification

a. Indexing Files

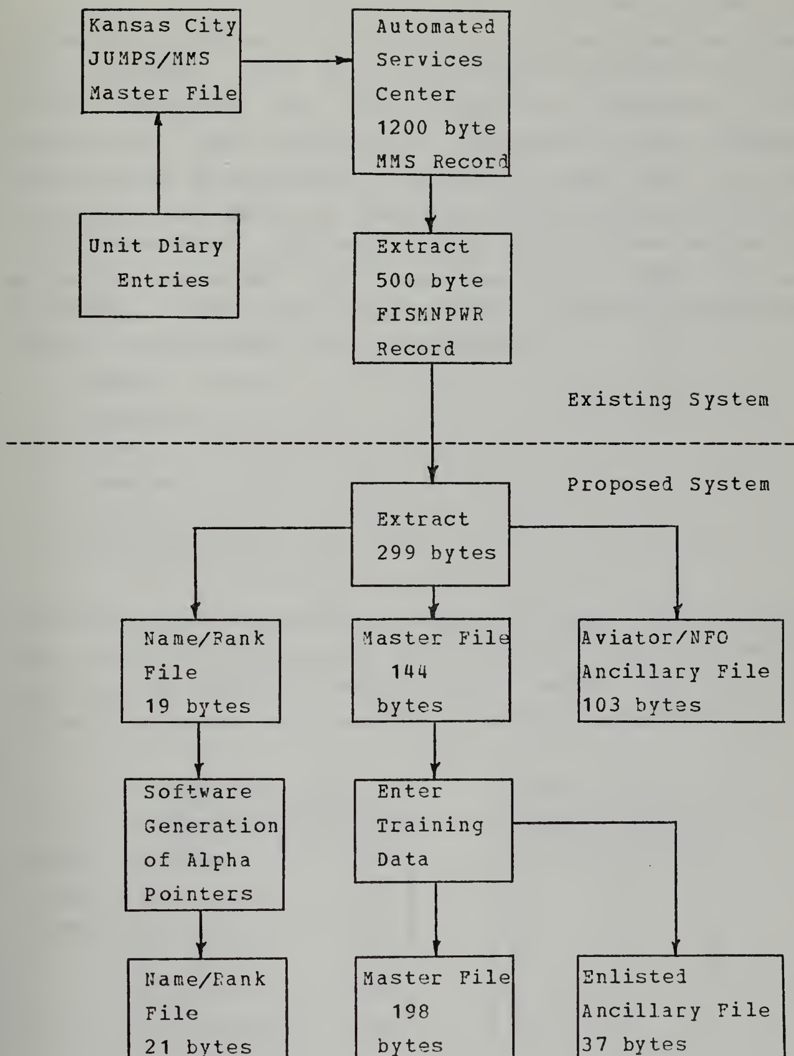
The NAME/RANK and Master Files can both be indexed by the Data Base Number (DBN). The introduction of DBN to identify an individual's Master record has two advantages over using name or social security number. First, it provides a direct index into the data base. No searching or hashing is required. Second, it is a faster entry for the operator in that at most four characters are required to identify an individual. This could be time saving when the records of 50 or more individuals must be updated. Ancillary files, however, do not contain the same number of records as the Master file. Thus, it is necessary to index these files with different indices. A Master Directory can be used to achieve the translation from DBN to the index number for the desired ancillary file. Thus, if the record from the second ancillary file is required, the second field in the Master Directory, indexed by DBN, would

specify the index into the ancillary file. The linkage between ancillary file name and its field in the Master Directory is discussed in the section on operator communications.

b. Establishing Master Records

Master File records may be established by reading the header information from the FISMNPWR File. If a record already exists, it is updated; if not, then a new record is established. Whenever a new master record is established, the individual's name must be inserted alphabetically among names in the NAME/RANK File. An index is assigned to the record. This index is retained for the life of the record in the system and can then be used as a reference number, hereafter referred to as Data Base Number (DBN).

The initial creation of the data base will be accomplished by creating individual headers for the master file. Since all header information exists on the FISMNPWR file, this file will be used as the source data for creation. The local data elements will be added at this time, or as is convenient for the user. The important point is that all individual records will have been created. The following flowchart depicts the initial creation of local files.



Initial Creation of Local Files

Figure 1.

As a sample of what is accomplished, the following narrative is presented. At file creation time, the FISMNPWR file is sorted alphabetically by last name within reporting unit code (RUC) or equivalent unit designator. The information contained in each FISMNPWR record which is pertinent to the local data base is then extracted and used to form the header information of the local Name/Rank File and Master File. The system is primarily concerned with the last name field of each individual. For example, suppose the following individual records in RUC 99999 are to be created.

Adams..........
Johnson........
Jones..........
Jenkins........
Smith..........
Williams.......

Since all files are empty at creation time, these records will be added sequentially to the files. Data base numbers will also be assigned sequentially. After creation, the files in question will have the following appearance:

Name/Rank File

		Name ///	Alpha Ptr
Data	1	Adams	2
Base	2	Johnson	3
Number	3	Jones	4
	4	Jenkins	5
	5	Smith	6
	6	Williams	$

Master File

	Common Data
1	Header/Trailer
2	
3	
4	
5	
6	

c. Establishing Ancillary Records

Ancillary records may be established for a variety of data. They may be established automatically when data from FISMNPWR is input or manually by operator action. A simple file management scheme for ancillary records is to maintain a parallelism between the indices of the master record and the ancillary record. This requires allocation of storage for the same number of ancillary records as master records. Such an approach is efficient when a specific ancillary record is likely to be created for a large percentage of the master records. However, the purpose of ancillary record design is to permit creation of different types of records for individuals. Some ancillary records may be created for as little as 10 percent of the master records. Also, as the system evolves, new ancillary record types may be desirable. Thus, to use auxilliary storage more efficiently, the ancillary records should be chained to the master through pointers. This chaining requires applications programs to update pointers from the master to the ancillary record. For each ancillary file which is established, the amount of storage available for ancillary files must be decremented by a free storage management program. Ancillary records may be deleted individually. When a record is deleted, pointers are adjusted and the storage which is made available is added to the free storage.

d. Master Directory

If the individual is to have an ancillary file created from FISMNPWR data (i.e. Aviator/NFO File), then an appropriate entry will be made in the master directory pointing to the record in the file. The FISMNPWR input record (local header) must allow for all possible data that are to be entered into the local data base. The Name/Rank

file, the Master File and the Aviator/NFO File are all created from the FISMNPWR record. Therefore, the FISMNPWR input record is 267 characters long.

Once the data base has been created, the common transactions of add, change or delete may be processed against it. Only the ancillary files may be added or deleted at the local level. The creation of ancillary records is based on the fact that an individual possesses a Master File record. Deletions may be made to these ancillary files at any time.

e. Transactions

Changes are the only transactions allowed at the local level to the Name/Rank File, Master File, or Aviator/NFO File, and then only to specified fields of the records. In fact, the operator should only be changing trailer information in the Master File. The reason for this is to maintain data security and integrity. Although duplicate copies of data elements may exist, it is important that source data changes to these elements be accomplished at only one site.

The FISMNPWR data base is recreated on a weekly basis. This is because the transaction rate is low and the unit diary reporting system is slow. For the most part, the FISMNPWR file is simply a selection of data extracted from the MMS file. The MMS file transactions are processed in only one location, Kansas City. Thus, satellite installations would not process transactions against the file. The weekly creation of the file reflects transactions which have occurred over the last seven days which have been processed in Kansas City.

The training data base derived from the

FISMNPWR file will also reflect transactions which have occurred at Kansas City through the unit diary reporting system. Due to the small transaction rate anticipated at the local level, it is deemed appropriate that the training data base need not be updated more than once per week. The process of adding and deleting records at the local level is addressed to show the tentative processing required.

Suppose that the FISMNPWR file is sorted and delivered to the local computer for processing with the following names:

Adams
Jones
Jenkins
Smith
White
Williams

This reflects one addition, White, and one deletion, Johnson, from the previous file.

The addition and deletion (update) process may be viewed as a match/merge process. Since the transactions have previously occurred for header information, no method of determining what the transactions were exists. The incoming file from FISMNPWR is matched against the existing local file. If a match occurs, the input record, reflecting changes, will be copied to the existing record. If the next input record does not match, the collating sequence will indicate an addition or deletion. If a record is to be deleted, pointers are adjusted and the entire record pertaining to an individual in each file is erased. A list will be kept of the open slots in the data base numbers list that occur because of deletions. For example, after Johnson has been deleted the files have the partial appearance:

Name/Rank File

Data Base Number	Name /	Alpha Ptr
1	Adams	2
2		
3	Jones	4
4	Jenkins	5
5	Smith	6

Master File

	Common Data
1	Header/Trailer
2	
3	Header/Trailer
4	
5	

Thus, slot 2 in the data base number list is left open. Jones, Jenkins and Smith match, so only their header information is copied. If a record is to be added, the first open slot in the data base index is found. Pointers are then adjusted and the new record is added. For example, after White is added, the files look like:

Name/Rank File

Data Base Number	Name ///	Alpha Ptr
1	Adams	2
2	White	6
3	Jones	4
4	Jenkins	5
5	Smith	6
6	Williams	$

Master File

	Common Data
1	Header/Trailer
2	
3	
4	
5	
6	

At the completion of the process, all additions and deletions should be printed for audit and a revised list of data base numbers should be printed. Note that once a individual record is created, the data base number is kept throughout the life of that record.

The process which has been described aids in the updating of files by conserving storage space and maintaining a data base which ensures fast retrieval. A reasonably simple backup system, to be discussed

subsequently, may also be implemented.

f. Backup

A disk to tape copy utility may be used to create a copy of the existing data base before update. This can be done weekly such that a father tape file is maintained with the son (current file residing on disk). The tape file, together with the transaction tape kept for audit trail purposes, gives suitable backup at the local level. Additional backup capability is provided for the FISMNPWR data base. Because of the relatively small size of the local training data base, re-creation of the FISMNPWR data elements for local backup would not be difficult. However, the local facility needs the backup because of locally generated data.

B. FILE MANAGEMENT SOFTWARE

1. File Structure

The files in the system consist of:

1. full lenth files such as the Master, Name/Rank and Master Directory
2. ancillary files whose lenth is less than or equal to the full length files.

The number of records in full length files and pre-defined ancillary files such as Aviator/NFO is a compile time parameter which is dependent upon the type of unit. The number of records in operator-defined ancillary files is specified at run time, giving rise to variations in file length among ancillary files. Full length files are all of the same length. All records are fixed length. The number of records in operator-defined ancillary files is assigned by the system when the file is defined. The number of

records is chosen to be a convenient physical portion of the auxiliary storage medium. When this space is exhausted, the system must provide for the linking to new areas or for extending the area by compacting existing files.

2. Alphabetizing Records

In order to facilitate the alphabetical listing of names and location of record by name, the NAME/RANK file should be maintained in alphabetical order. Although this process is costly in processing time the justification lies in the assumption that the number of records in the file will be small and additions and deletions will be infrequent. Physically arranging records in alphabetical order would make location of names easy but it would require moving large numbers of records in auxilliary storage. This shuffling would cause DBN's to change frequently, thus confusing the system user. The alphabetical order should be maintained through a sequence of pointers which can be used to chain through the indices of the records in alphabetical order. To prevent chaining through the entire structure when adding a new record, a table containing some number of alphabetical partitions should be defined. The partitions should be selected to provide equal distribution of names among the partitions. Partitions are defined by the first two letters of the last name. Each partition contains the DBN of the first and last name within the partition. The partition table can then be searched binarily. For example, if there were 64 partitions each of which is expected to contain 10 names then at most 6 comparisons would be required to locate the partition and 10 record accesses to locate the record within the partition.

3. Access of Files

The small expected number of records per file, the relatively slow system response time requirement and the low frequency of output requests indicate that direct access of system files is not absolutely required. Although a sequential access method meets all the requirements stated in this thesis for a training information system, it does not provide for system growth. The training information system, viewed as the forerunner of a general purpose information system, incorporating many squadron/battalion functions, should be implemented on a direct access storage device (DASD). The software descriptions in this section are predicated upon this approach.

4. Retrieval of Records

Records may be retrieved for input of new information or for compiling lists for output. In the former case the primary method of specifying a record is data base number. This permits direct retieval of the master record. If an ancillary record is desired, the pointer in the Master Directory is used for direct retieval. Often, retrieval of ancillary records for several individuals is required. This requires two accesses of auxiliary storage for each individual to retrieve, first, the Master Directory item and then the ancillary record. Updating can be accomplished more efficiently if the Master Directory is core resident during this processing. When retrieval by name is used, the alphabetizing procedure can be used to obtain the DBN. Specifying ancillary records by name of the individual has the following disadvantages:

1. A search is required to locate the DBN.

2. Two accesses per ancillary record retrieval are required unless both the name file and master directory are simultaneously core resident. Additionally, there is the

inconvenience of entering more characters to specify a name than the four required for data base number. In compiling lists for output, a block of records can be retrieved and checked.

5. Addition and Deletion of Records

Since all files consist of a fixed number of fixed length records there are two alternatives for managing the addition and deletion of records.

1. Assign the first available storage location in the file when a record is added. When a record is deleted do not alter the remainder of the file and insert an availability bit. This method requires a bit in each record to indicate that it is either in use or available. Further, this bit must be checked each time a record is accessed to ensure that the record contains valid information. The advantage of this method is that pointer updating is simplified since records are never moved.

2. Assign the first available storage location in the file when a record is added. Delete a record by replacing it with the last record in the file. This method, while eliminating the need to designate gaps in the file, requires additional processing to update the system of pointers in the Name/Rank file and Master Directory. Additionally, an intermediate table which translates DBN to Master file index must be maintained if individuals are to retain the same DBN.

Because of the search method to be used, the first alternative was selected. The checking of the "in use" indication can be incorporated into the masking used during searching.

6. Searching of Files

The primary consideration in determining how to search records is whether to link the records with the same attribute values. This could be accomplished by threading the records with pointers or creating inverted files. This level of sophistication is not warranted for the following reasons:

1. The number of records per file will always be small. This means that the worst-case search of all records is bounded because the number of personnel in the unit determines the number of records. The efficiency of a sequential search will diminish only slightly as new functions are added to the system. The small increase can be attributed to greater record length.

2. The search time may be dominated by the record access time. If the entire file to be searched is core resident, then the search can proceed more quickly with linked attribute values. If, however, only a portion of the file is in core, then the linking of attribute values may cause more record retrieval from auxiliary storage by requiring the access of new blocks of records each time a new pointer is processed. The time required to retrieve a record is large compared to the time required to search it.

3. It is difficult to determine which attribute values should be linked. Only frequent retrieval requests for a particular attribute value would compensate for the additional processing required to manage the linked lists. It is not known if any special retrieval requests will recur with regularity.

Therefore, the search should be of contiguous records within a file. The probem is that a number of

attributes may be checked within each record. One solution is to make multiple passes through a file checking for one attribute value on each pass. A better solution is to prepare a mask based upon the structure of the attribute values which are to be searched. A block of records from the file can then be brought into core and searched by performing logical masking operations on each record. When a "find" occurs in the master file, the print bit can be set in the NAME/RANK file to provide an indication that the name has been selected for printing. The same technique can be used when searching the ancillary files. In this case the reverse ponters to the name file must be used in order to set the print bit. In order to minimize the number of disk accesses the NAME/RANK file should be in core when this searching takes place.

There is additional processing when two or more files must be searched. Multiple masks must be created and additional logic must be implemented to set and clear the print bit in the NAME/RANK file.

7. Scheduling

Once the search has been completed, the indication of all names which have met the search criteria is left in the NAME/RANK file. If the system user desires to further reduce this number, scheduling programs can then be applied to this list of names to produce the reduced list for output. The process involves counting the number of elements in each category and reducing that number by some proportion. The example below illustrates this. Suppose 20 rifle range quotas were to be filled in such a way that no section bore a disproportionate burden. The user could specify a search of the data base for all E-3 and below who had not fired the rifle during this calendar year. To this result he could choose to apply a sheduling algorithm as

follows:

SCHEDULING OPTION: Proportionately by rank and section. Select 20.

After the search of all records the following numbers of records were found to have the requisite values.

NUMBER	SECTION	RANK
7	1	E-2
4	1	E-3
7	2	E-2
8	2	E-3
2	3	E-2

The scheduling algorithm is then required to select 20 from these. The result would be as follows:

NUMBER SELECTED	SECTION	RANK
5	1	E-2
3	1	E-3
5	2	E-2
6	2	E-3
1	3	E-2

The actual printout is the list of names of individuals selected for the rifle range detail. The number of scheduling routines depends upon the desires of the user. Additional programs can be added at a later date. This thesis considers the inclusion of only the following schedulers.

1. Proportionately by rank and section

2. Specific numbers by rank, section or rank and section.

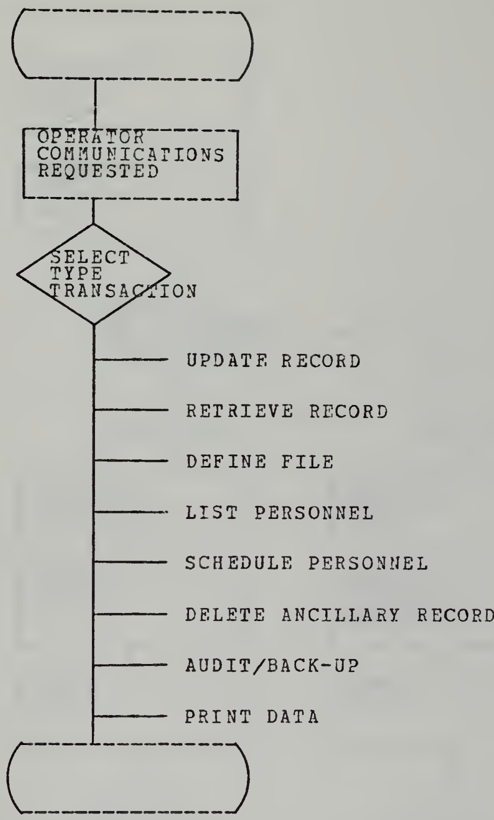
Additionally, sections may request that certain individuals not be scheduled. There must be a capability to exclude individuals from consideration by the scheduling algorithm.

8. Operator Communications

A retrieval language is required to provide a means of operator interaction with the system. The method of operator communications herein described is semi-conversational with rigidly structured responses. This language should not pose an obstacle to system use because relatively few responses must be learned. The operator is required to input data description words and numbers, and predefined command words which select a major type of operator request processing. If the system were expanded to perform other functions, each additional function would require an addition to the language vocabulary but not to the complexity of the language structure.

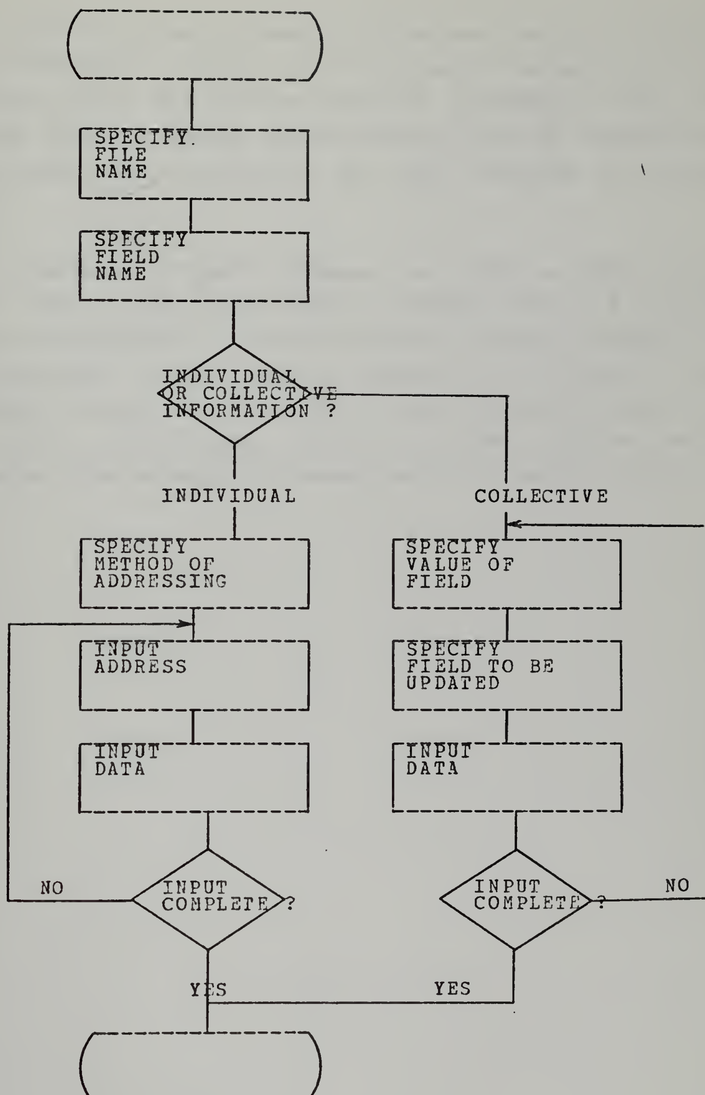
The operator communications processing requires programs to output pre-stored inquiries and input the operator response. Operators are required to input file and field names (attributes) as well as attribute values. Operator input field values are matched with internally stored names linked to actual physical locations. A file directory containing each file name, base address on auxilliary storage and a pointer to a list of field names for each record serves this purpose. The list of field names contains the name of each field, the location of the field within the record, the type of the field (numerical or character), the type of access, and the range of values for which the field is defined.

The following flow charts depict the operator interractions required to initiate system processing tasks. Explanations of the specific responses follow the flow charts.



Selection of Major Processing Task

Figure 2.

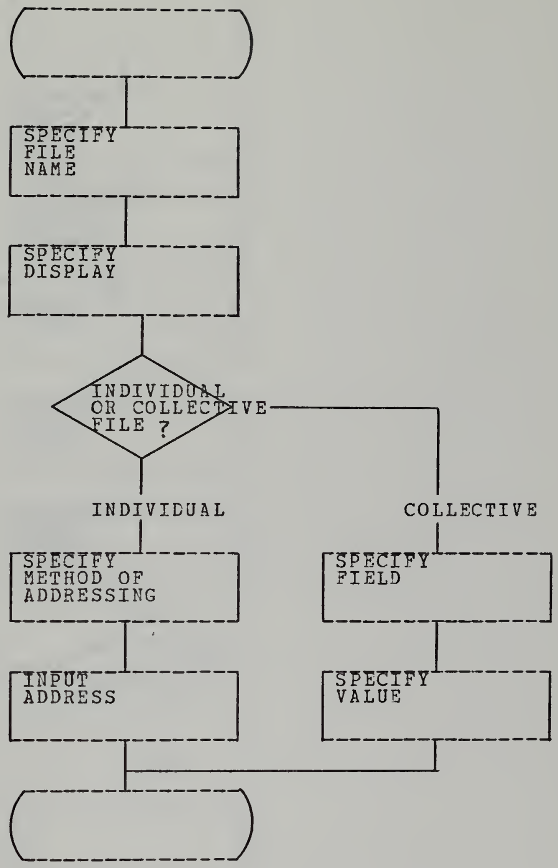


Update Record

Figure 3.

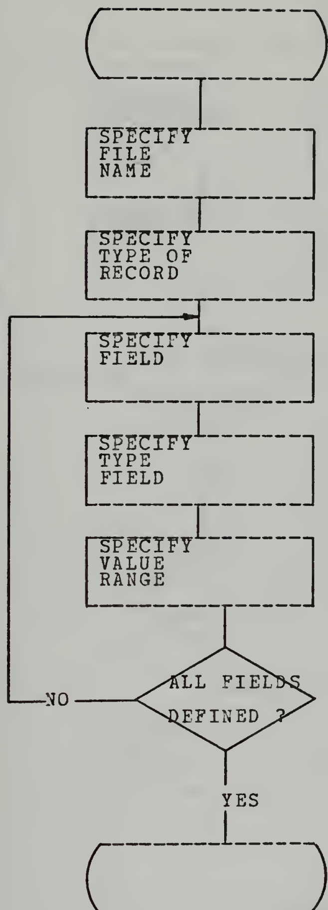
The input can be used to access and update records as it is entered. However, since we need to record all data for later audits, the updates could be performed from the recorded data. A header describing the type of transaction and the date should accompany the data entries for each record.

Whether the input address is NAME or DATA BASE NUMBER (DBN), the system should retrieve both and display them for the operator to ensure that the correct record is being addressed. When a name is entered, it is found in the NAME/RANK file and then the DBN is retrieved and written to the input audit tape. Input by NAME requires more processing time as well as a more lengthly operator entry.



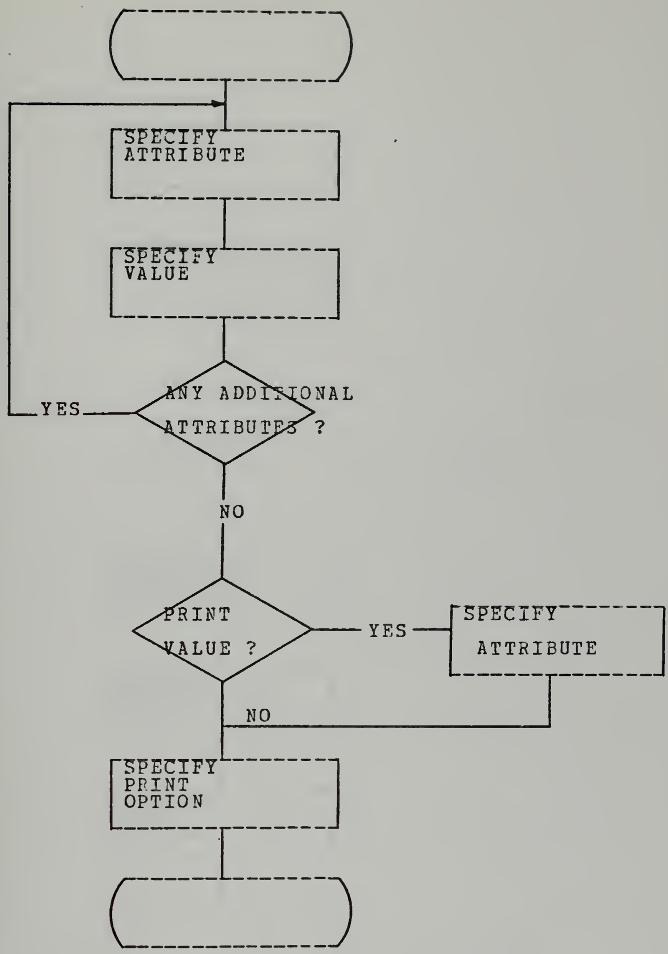
Retrieve Record

Figure 4.



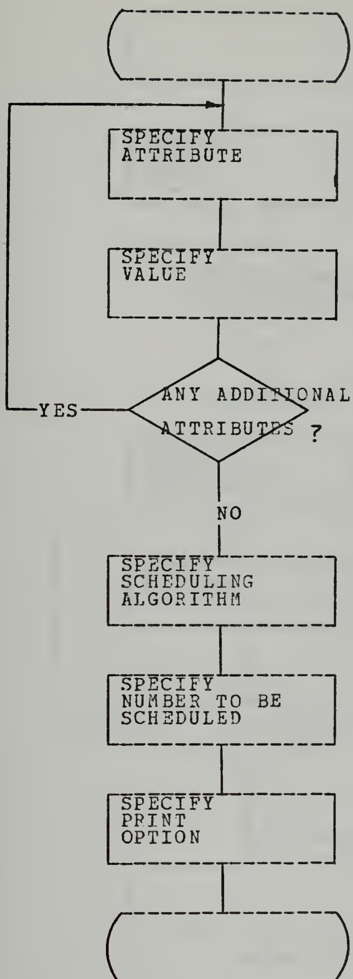
Define File

Figure 5.



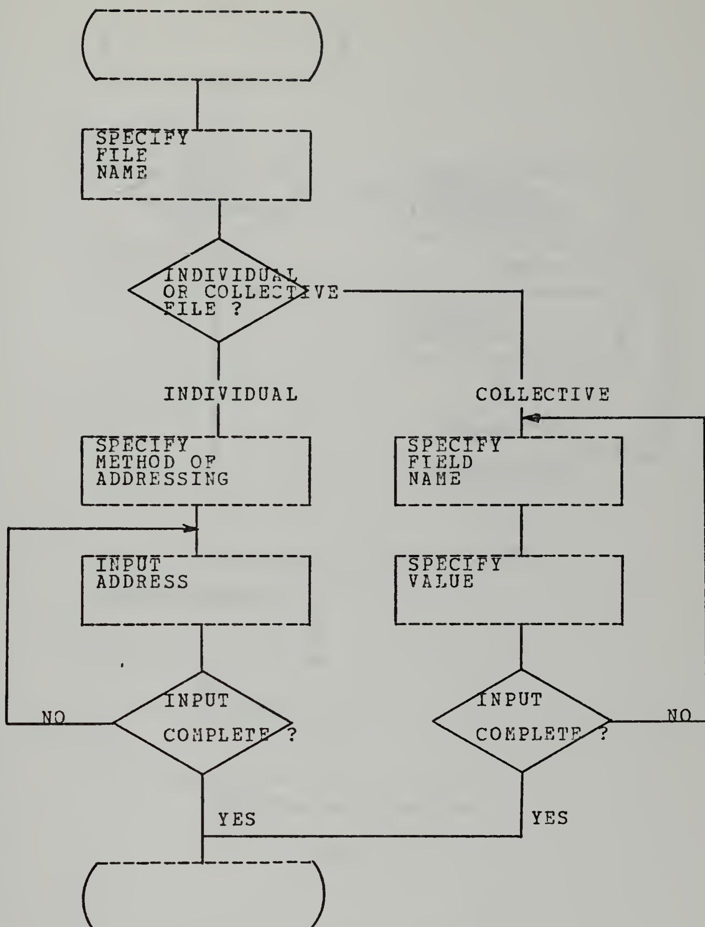
List Personnel

Figure 6.



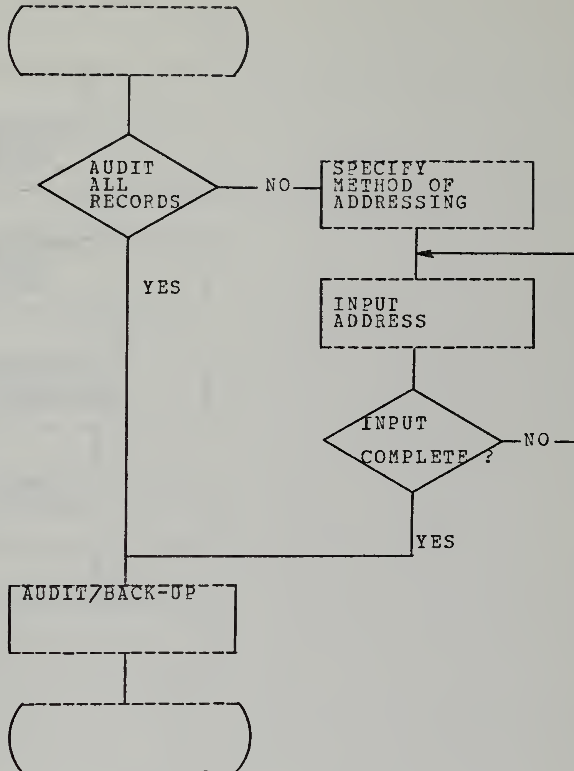
Scheduling Personnel

Figure 7.



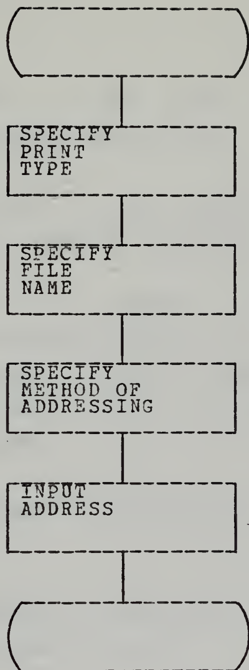
Delete Ancillary Record

Figure 8.



Audit/Backup Creation

Figure 9.



Print Data

Figure 10.

OPERATOR RESPONSES:

SPECIFY FILE - The exact name of any file in the File Directory must be specified.

SPECIFY FIELD - The exact name of any field within the selected file must be specified.

INDIVIDUAL OR SUMMARY - Individual records contain information to individuals and indexed by DBN or NAME. Summary records contain information pertaining to types of training. Summary records are indexed by the value of the field specified.

SPECIFY METHOD OF ADDRESSING - Once the summary record itself has

SPECIFY METHOD OF ADDRESSING - Individual records may be addressed by DBN or NAME.

SPECIFY FIELD TO BE UPDATED - Once the summary record itself has been identified, the field to be updated must be specified.

INPUT COMPLETE - The operator continues to input record address and then data until he indicates that the input list has been completed.

SPECIFY DISPLAY - The record may be displayed on the CRT, printed or both.

SPECIFY TYPE RECORD - If the record is an individual record then provision must be made for linkage through the Master Directory by pointers.

SPECIFY TYPE FIELD - Fields may be character or numerical.

SPECIFY VALUE RANGE - This is the valid range of values over which the field is defined.

SPECIFY ATTRIBUTE - An attribute is an exact field name within any file.

SPECIFY VALUE - This is the field value for which the search is to be conducted.

ANY ADDITIONAL VALUES - There is a capability to " and " values. The operator may enter additional attributes and values for this purpose.

PRINT VALUE - If a field value is to be printed with each name then it can be specified here.

SPECIFY PRINT OPTION - The operator must specify one of the system listing options.

SPECIFY SCHEDULING ALGORITHM - The operator must specify one of the system scheduling algorithms. He then specifies the number to be selected.

ALL RECORDS - If ALL is specified then all records are audited. If not then the addresses of specific records must be furnished.

AUDIT/BACK-UP - If AUDIT is specified then results 1re printed. If BACK-UP is specified then back-up records are created.

SPECIFY PRINT TYPE - The operator may choose to print a system table or an individual record.

9. Validity Checks

Each attribute is defined over a specific range of values. With each stored field name the range of values and the type of field, numerical or character is stored. When a field is updated a program must check that the input is of the prescribed form and within the range of values. An indication should be provided to the operator whenever an attempt to input invalid data is made. The invalid entry is then displayed for the operator along with an indication of why the entry is invalid.

Read access is granted for every file and field contained in the File Directory. Write access will be permitted only for training data. Write access is denied all fields updated by FISMNPWR files.

10. Input Restrictions

The names of files and fields within records must be restricted in length. All file names, field names and individual names must be specified exactly. There will be no attempt to correct an erroneous input by juxtaposition of the various letters.

11. Audit

The method of conducting audits is a procedure established by users. This discussion considers that audits are conducted for two purposes: to verify that the internal system functioning is correct and to ascertain that the information itself is correct.

The system verification processing occurs prior to the creation of each new back-up record. At any given time

there exists within the system a dated back-up record and an input tape of all updates to the record since that date. This provides a capability to recreate the current data base. The input tape contains date, DBN, a code specifying the file and field to be updated, and the input data. The verification consists of scanning the input tape and updating a copy of the back-up record. If the result is not the same as the current record then an error indication is printed. If it is the same, then the current can become the file back-up.

When the contents of a record is to be audited, the process is the same as above except that they following information is printed:

1. Current Record
2. Number of differences between Current Record and updated back-up record
3. Back-up Record
4. Updated Back-up Record

Items 3 and 4 are only printed when differences exist. The individual is then asked to verify the contents. Disputes can be settled by hard copy rosters of input data. Audits may be conducted for individual records or for all records.

12. <u>Definition of Files</u>

To satisfy the requirement for purely local data storage and record keeping, ancillary files which can be specifically defined by the user are available. The number of ancillary records which can be defined is limited only by the following.

1. The number of file and field names which can be placed into the File Directory.

2. The available storage on the DASD for the

physical location of the records.

The number of ancillary records which can be created for an individual is limited by the number of pointers which can be placed in the Master Directory. As the operator defines each field for a new ancillary file, the names are placed into the file directory with the valid range of values for each field.

The field name, size and type are specified and stored in the file directory. Each newly defined ancillary record is referenced through a pointer in the Master Directory. The field in the Master Directory in which the pointer can be found is entered in the File Directory.

Files composed of records not representing individual marines can be defined. These collective files cannot be indexed by DBN. In this case a field name and value is used to identify a desired record. A collective file can be used to satisfy the requirement to retain information concerning subjects taught and numbers attending.

The following diagrams illustrate the files discussed for the system.

MASTER FILE

Indexed by DBN

Contains data common to all marines

MASTER DIRECTORY

Indexed by DBN

Each record in the Master Directory contains pointers to the ancillary records defined for that individual.

Fields within the record denote the specific files. Fields are referenced by ancillary file number

Pointer to record in file 1.	pointer to record in file 2.					

Master File and Master Directory

Figure 11.

FILE DIRECTORY

Indexed by ancillary file number		
Indexed	1	Name of ancillary file 1
by	2	Name of ancillary file 2
ancillary	3	
file	4	
number	5	

ANCILLARY FILES

Indexed by pointer contained in the Master Directory	Each ancillary file contains data relevant to a subset of marines

File Directory and Ancillary Files

Figure 12.

13. Print Formatting

Information to be printed can be subdivided into the following categories:

1. Pre-defined text

2. Records

3. Name lists

4. Special system data

Pre-defined text consists of all header information. Several coded headers ranging from titles to column headers for lists can be stored on auxiliary storage and printed in conjunction with other information. Records are printed by obtaining the text for each field name from the File Directory and printing it with the value of the field obtained from the record itself.

The printing of lists of names is the consequence of many different operator requests. In the processing of lists of names it is necessary to chain through the names in the NAME/RANK file from a to z and check each to see if it was selected by the search programs. Multiple passes through the NAME/RANK file are necessary when the data is to be sorted by rank and/or section. This procedure can be justified by the fact that few records need be examined on each pass and that system response time does not appear critical. When a retrieval value is to be printed with the name, it is stored in a VALUE FILE paralleling the NAME/RANK file. File definition which the user may desire is printed, such as a list of DBN'S or the File Directory.

All print data is formatted and written to DISK before being output. Procedures to convert binary data to characters are necessary.

C. OPERATING SYSTEM

The Operating System provides primarily the interface between the system and the peripheral devices.

1. CRT

The Operating System (OS) must interpret interrupts to determine that the user desires to interract with the system or has completed interaction with the system. It must handle all input/output buffering. For input the OS establishes one line buffers, the contents of which are examined by the Operator Communications program. Full buffers are passed to the OS for output by the operator communications programs or as a result of the retrieval of an individual training record.

2. Printer

The OS must provide for the control of output to the printer. The PRINT procedures build the output and store it on the DASD. The OS must then pack the data into output buffers and provide for output buffering.

3. Tape

The OS must provide for the control of input/output to the tape drives. Output is built by the calling procedures. It must provide for the reading of blocks of data from tape, for purposes of updating files and conducting audits.

4. DASD

The OS must provide for the mapping of logical records to physical storage. It must therefore maintain a table which contains the location of each file. When a file is defined by the user the OS must obtain the free storage, if it is available, and assign the data to cylinders on the disk. It must maintain tables showing the spindles, cylinders and tracks on which the data can be found. Periodically the OS must move existing files in order to consolidate fragmented free storage.

When data is retrieved the OS must be able to locate on the disk the correct record when a file name and index is specified by the calling program. For searching, the OS must be able to retrieve a block of records. This block is then searched by the calling procedure.

Additionally, the OS must provide for the storage of applications programs on the DASD. This requires a directory of program segments and their locations. Each module is capable of calling another module through the OS.

Finally, the OS must provide for the spooling to Disk of the print information. It must maintain a queue for information to be printed.

D. INTERACTION OF PROCESSING TASKS

There appears to be no requiremnt for multiprogramming. The light processing load and the absence of rigid turnaround constraints on the processing tasks enable use of a first-in-first-out priority system. Provision must be made, however, to overlap input/output tasks and internal

processing. Specifially, the printing task is the most time consuming. It may be desirable to input and process a request while the previous one is printing out. Provision should be made to queue the printing of information. The print portion of the OS should be initiated by the applications prgram and performed concurrently with processing and input.

The various processing tasks can be segmented into phases of transaction. The operator input phase continues until a complete request has been interpreted and any input data have been spooled to tape. At this point either a retrieval for update of records, a search, or some special processing such as listing DBN's or conducting an audit, is queued. Any output from this phase is written on the DASD. If names are to be listed, the print bits in the Name/Rank file are set prior to formatting the print data on DASD. The final step is the queueing of output data for printing. Table II divides the processing tasks into phases. The programs and data base are listed with the estimate of the storage requirement of each. If that program or data is core resident during a task, a one is placed in the column for that task.

PROGRAMS/DATA BASE	OPERATOR COMMUNICATIONS	INPUT OF DATA	RETRIEVAL OF RECORDS	SEARCHING OF RECORDS	DEFINITION OF FILE	SCHEDULING OF PERSONNEL	AUDIT	PRINT FORMATTING *SYSTEM DATA*	PRINT FORMATTING *NAME/VALUES*	READING OF FISMNPWR DATA	DELETION OF ANCILLARY RECORD	COMPACTING OF FREE SPACE	UPDATING/CREATION OF RECORDS	PRINTING OF INFORMATION	CRT DISPLAY OF INFORMATION	CREATION OF BACKUPS	CORE
OPERATING SYSTEM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		6K
STORAGE HANDLER					1							1					1
OPERATOR COMM MODULE	1													1			4
INPUT TO AUDIT TAPE		1												1			.5
AUDIT TAPE TO DATA BASE										1			1	1			.5
SEARCH ROUTINES				1										1			1.5
FILE DIRECTORY BUILDER					1									1			.5
MASTER DIRECTORY LINKER					1						1			1			.5
SCHEDULE ALGORITHM						1								1			.5
AUDIT CHECKER							1							1		1	1
PRINT FORMAT ROUTINES			1					1	1					1			1.5
NAME/RANK FILE		1	1	1		1			1	1			1	1			14
MASTER FILE				1		1				1			1			1	130
ANCILLARY FILE(S)				1		1										1	25
VALUE FILE									1								13
MASTER DIRECTORY					1						1						75
FILE DIRECTORY				1	1												.5
PROGRAM DIRECTORY	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	.5

Core Requirements

Table II.

By inspecting Table II it can be seen that the critical internal memory requirement occurs when one request is being processed while another is being initiated. A total requirement of 31.5K 16 bit words exists. Therefore, a minimum of 32K words would be required for the system.

E. SIZE OF THE DATA BASE

The size of the training information data base is dependent on the size of the record of each of the files and the number of individual Marines in a squadron/battalion. The length of each of the records has previously been described. After examining Marine Corps Tables of Organization (T/O's) the following sizes of units were determined:

UNIT	NO. OF T/O'S INVESTIGATED	AVG SIZE OF UNIT	MAX SIZE OF UNIT	MIN SIZE OF UNIT	
BATTALION	15	739	1149	406	TOT
		45	157	28	OFF
		701	1103	414	ENL
SQUADRON	14	358	697	162	TOT
		41	65	18	OFF
		317	635	128	ENL

Assuming that the system should be sized by using the maximum figures from this table and allowing for extra storage due to possible additions to avoid overflow, the following table defines the size of each of the existing files in the data base.

FILE	RECORD LENGTH	NUMBER OF RECORDS	SIZE OF FILE
NAME/RANK	21	1250	26,750 chars
MASTER	209	1250	261,250 chars
AVIATOR/NFO	103	75	7,725 chars
ENLISTED	36	1200	43,200 chars
		TOTAL	338,925 chars ***

Although this table shows the maximum size of the data base, without all ancillary files, the size of the data base at each Squadron/Battalion would be different and would be redefined at system generation time in each unit. The maximum figures are used to determine hardware characteristics which must be met.

IV. IMPLEMENTATION CONSIDERATIONS

A. ALTERNATIVE METHODS

Once the requirements of the desired information system have been defined, the developer faces several options concerning the detailed design, programming, and implementation of the system. He may choose to use an off-the-shelf system; he may contract with a systems house to develop the required system; or, he may design and program the system with in-house personnel. In order to make the correct decision, the costs associated with all three available courses must be estimated. This includes not only the dollar costs of systems acquisition and personnel training but also the time and operational effeciencies and inefficiencies associated with each method. The costs of maintaining the developed system must also be estimated. When all costs are tallied a decision as to the feasibility of the proposed system should be made.

B. GENERAL COMPARISON OF METHODS

Perhaps the greatest advantage inherent in using an off-the-shelf system is that the user can avoid the extensive programming, debugging, and implementation costs associated with the development of an original system. In addition, providing the system is sufficiently general, it may be much easier to meet additional applications needs throughout the system life.

The generality of an off-the-shelf system has several disadvantages. It often causes relatively long execution times and large memory requirements.

Several very significant advantages accrue from the

employment of a systems house for the design and implementation of the required information system. First, the systems house approach, when compared to an off-the-shelf system, leads to a system more nearly in accordance with the specific requirements defined in the development contract. Second, the system benefits from the expertise associated with the past efforts of the system house. Third, the contract may be written so that the system house is responsible for both development and maintenance of the information system or for development only. Hence, an organization may utilize this alternative because its resources are sufficient to maintain and modify an existing system but not to develop a system from scratch.

The main disadvantage associated with the systems house approach is its high cost as opposed to that of an off-the-shelf or self-developed information system. In addition, the user of a systems house may not realize the quality control and responsiveness that he would probably achieve if he utilized his own resources. Another possible problem associated with the systems house is a business failure in the midst of system development or maintenance.

The advantages related to a self-developed system are derived from an increased control over the design and implementation of the information system. The feedback loop involving design, debugging, and implementation is much tighter than with the other alternatives. If an organization possesses manpower resources of sufficient skill available for such a dedicated effort, it may realize considerable savings. Disadvantages are closely related. In order to acquire sufficient internal resources, an organization may have to prolong development time. In the worst case, an organization would be incapable of completing a system. The organization would have to rely on outside help under unfavorable conditions.

C. CRITERIA FOR DEVELOPING COSTS

The problems of estimating monetary and time costs is in no way trivial. Failure to meet deadlines and cost overruns are the normal rather than the exception. Efforts have been made to develop a methodology for estimating the costs associated with computer program development. In particulalar references 10 and 11 present qualitative and quanitative guidelines for estimating costs in terms of man-months, computer hours, and months elapsed.

The approach taken in reference 10 is to divide the development process into major activities which are further subdivided into phases and tasks. The three activities and their associated phases are:

1. Analysis and Design Activity
   - System Analysis Phase
   - System Design Phase
2. Program Implementation Activity
   - Program Development Phase
   - Program Coding Phase
   - Program Checkout Phase
3. Support and Turnover Activity
   - User Documentatation Phase
   - User Training and Assistance Phase
   - Turnover Phase.

The study then defines the tasks associated with each phase. Each task definition includes a description of the relevant cost factors along with a means to estimate the given costs.

While this format does not exactly describe the process required to cost the system proposed in this thesis, it does provide a sound basis for developing cost estimates. The Analysis and Design Activity should be broadened to include phases devoted to the selection of hardware and software. This selection then becomes the premise upon which the plans

and costs of subsequent activities are defined. As previously mentioned costs associated with the operation and maintenance of each system must also be developed. A set of resource costs associated with each combination of hardware and software will result. The resulting set of costs should be instrumental in selecting the proper system for development.

D. LANGUAGE CONSIDERATIONS

In the event that the developer elects to design and implement his own data base management system, he is immediately confronted with questions concerning the language to be used. A decision must be made whether to use a low level assembler language or a high level language.

The advantages and disadvantages of high level languages are treated by Sammet [Ref. 8]. They are summarized as follows:

1. Advantages of high level languages
    - Ease of learning
    - Ease of coding and understanding
    - Ease of debugging
    - Ease of maintaining and documentation
    - Ease of conversion
    - Reduced elapsed time for problem solving
2. Disadvantages of high level languages
    - Time required for compiling
    - Inefficient object code
    - Difficulties in debugging without learning machine language
    - Inability of the language to express all needed operations
    - Advertised advantages do not always exist.

The major programming effort will take place prior to system implementation. Any subsequent programming will be to correct, modify, or add to present applications. This programming will be conducted at a centralized activity; it will not be undertaken by the end-user. In evaluating the tradeoffs involved in language level selection this must be taken into consideration. The advantages associated with a high level language take on added importance in the the context of the manpower transience of the Marine Corps. Regarding disadvantages, the following can be said. Since the user will be executing pre-compiled programs, the time requirement for compilation is not a factor in the field system. Inefficient object code can be overcome by optimization techniques which would be relatively inexpensive because of the one time effort. The same line of reasoning applies to debugging difficulties. In addition, many of the problems connected with compilation on a small machine can be overcome by compiling on a larger machine. However, this requires that such a machine be provided from existing inventory or new acquisition. Based on the above, a high level language should be used for programming the proposed system.

Important considerations in the selection of a language are discussed in Sammet [Ref. 8].

1. The language must be suitable for the application. In this case the language must have facilities for data base creation, manipulation, and modification. It should allow for easy file and record definition, manipulation, and transfer. Data should be available in numeric, string, or bit form. It is essential that the language have on-line cpabilities.

2. The language should be available on the desired computer. This is desirable but may be too much to expect in view of the fact that data base type applications are

relatively new in the minicomputer field. If a particularly desirable language is not available, consideration should be given to developing one. Compiler development, however, is an expensive activity. The expense might be justifiable if the language was to be used throughout the Marine Corps. As mentioned previously such a compiler is not required to execute on the small machine. The compiler may generate code for the small machine much more efficiently on a larger machine.

3. The language should be capable of efficient implementation on the desired machine. Compilation inefficiencies will not necessarily rule out a language. However, other structural incompatabilities will effectively block the use of a language. As an example, a language whose procedures are recursive might not be efficiently implemented on a machine without hardware stack facilities.

With these criteria in mind, a preliminary survey of programming languages has been made. While this discussion is not all-inclusive as to the languages available on present-day minicomputers, it does include the widely available languages. Many minicomputer manufacturers and original equipment manufacturers have designed specific languages for their particular systems. Several of these possess very powerful data base management facilities integrated with specific operating systems. The cost/benefit trade-off associated with these should be considered to the greatest extent possible in deciding upon a particular computer for system implementation. Several of these systems are discussed subsequently.

FORTRAN IV and other FORTRAN versions are available on nearly all minicomputers. However, FORTRAN possesses neither the data structuring or the data manipulation capabilities required for the proposed system. Consequently, it should not be considered as a candidate

language.

COBOL was designed as a common language for data base processing. It is the most widely used language for such purposes. As a consequence it enjoys a higher degree of convertability from machine to machine than any other language. COBOL is rapidly becoming widely available on minicomputers because of government requirements. In addition, COBOL possesses a powerful set of functional data base capabilities while at the same time retaining a relatively simple form. However, it was designed to operate in a batch mode and consequently has little on-line capabilities. In view of the on-line nature of the proposed system, COBOL would present substantial difficulties if used for the system.

BASIC was developed at Dartmouth College in 1965 with primarily scientific and scholastic applications in mind. It possesses only limited data base capabilities with some on-line capabilities. Because of its narrow data base capabilities, BASIC would also be difficult to use for the proposed system.

ALGOL is a scientific language available on relatively few present day minicomputers. It does possess limited data base facilities in its "record" data structure. It is primarily a batch type language which suffers from a lack of output format capability. ALGOL does possess some very favorable characteristics. Its logical block structure lends itself to easy reading and writting. It possesses bit manipulation capabilities and very powerful procedural capabilities. However, because of its stated limitations, ALGOL would present difficulties if used.

PL/I was initially intended to be developed as an improvement to FORTRAN. However its designers, IBM and the

[illegible]

IBM sponsored SHARE Group, decided to eliminate many of the restrictive features of FORTRAN and to incorporate desirable features of ALGOL and COBOL. The result was a very powerful and comprehensive general purpose language. A particularly desirable characteristic of PL/I is the existence of subsets of the language devoted to particular applications. This simplifies the user's learning process in that he must learn only that portion of the language pertaining to his application. A compiler may be streamlined by removing those portions of PL/I not not required. In any case, PL/I provides extensive FORTRAN-type scientific capabilities, data structures equal to those of COBOL, and Algol-type block structures and procedures. It has on-line capabilities and has been used in systems programming applications efficiently. The major drawback of PL/I is that it has been implemented on very few minicomputers. Its extensive capabilities necessitate a large compilation system. Thus it would probably be necessary to develop a compiler if PL/I were chosen. As mentioned previously compilation would be a centralized effort, not necessarily on the machine intended for field implementation. In this way the disadvantages associated with compilation could be overcome. As a consequence of PL/I's powerful capabilities it should be given consideration for system implementation.

E. HARDWARE/SOFTWARE ALTERNATIVES

It is beyond the scope of this thesis to recommend a particular hardware/software system for implementation. As an alternative, three systems are presented, ranging from a hardware only system, requiring complete OS and data base system development, to a turnkey system requiring a minimum of systems programming. As an example of a hardware only system, a Digital Equipment Corporation PDP-11/40 was chosen. The PDP-11 is the most widely used of its type

today. As such, many of its features are typical of present day minicomputers. An example of hardware/software system witha data base management and operating system, a Varian Data Machines V72, is presented. An example of a turnkey system, with a complete set of functions for the non-programmer user, a Microdata REALITY System , is presented.

The machine configurations presented do not constitute a recommended configuration for the training system. The hardware configurations for the Varian and Microdata systems represent the minimum core and disk resources which are required to support software.

A general system description and configuration for each system is presented in this section. For a detailed description of the system capabilities see Appendix D. The majority of the information on the systems was extracted from Datapro Reports on Minicomputers [Ref. 2].

1. PDP-11

Digital Equipment Corporation's PDP-11 Series is one of the most widely used minicomputer systems. The PDP-11 offers useful features in processor capabilities, instruction set capabilities and the UNIBUS.

The processor allows operands to be referenced in one of eight modes allowing operation as a stack machine, general-register processor, or memory-to-memory processor. Register architecture permits absolute addressing to all of memory, immediate operand specification, relative-direct addressing, and relative indirect addressing. Register architecture also facilitates table processing schemes.

The UNIBUS connects all system components and

peripherals. It is a single, bi-directional, 56 line high speed bus. The UNIBUS treats all devices the same in that the communication form between any devices is the same. This simplifies data manipulation while allowing devices to operate at the maximum rated speed.

The PDP-11 instruction set possesses many operands which can act on either bytes or words. This, along with available addressing modes, permits powerful byte manipulation capabilities.

The PDP-11/40 incorporates a memory management unit which provides additional main memory up to 128K and memory protect features. In many PDP Series machines the upper 4K of main memory is reserved for device addresses in connection with UNIBUS operation.

The following configuration is presented for comparison with the other systems. price information was derived from DATAPRO[Ref. 2]. It is approximate. Software delivered with the system is of the stand alone paper tape variety. Included is a stand alone assembler, editor, linking loader, on-line debugger and a single user BASIC interpreter.

Device	Pur price	Mo. maint.
CPU - PDP-11/40		
32K words	27,100	235
Disk cartridge system - 2.4 M Bytes	11,000	60
Tape drive and Control	10,745	95
Medium speed printer	17,500	75
CRT Terminals (4)	11,180	08
Total	77,525	473

Detailed information on system capabilities is contained in

Appendix D.

2. Varian 70 Series

The Varian 70 series was initially marketed in mid 1972. Distinguishable features of this machine include writable control store (WCS), dual port memory modules, and a memory map function. The most significant feature of the V 70 series in relation to this thesis is the TOTAL data base management system. base management system. Varian 70 series processors have a powerful microinstruction capability in the form of WCS. The WCS may contain user written reentrant segments of apllication programs, operating system, or user written extensions and modifications to the instruction set. Each microinstruction is 64 bits long and is individually addressable. Cycle time per instruction is 190 nano-seconds. WCS comes in modules of 512 word ROM with a maximum of four modules per processor. Micro-programming requires considerable systems knowledge and programming finesse. Errors in micro-code are difficult to locate and correct.

Dual port memory permits two units to access memory at the same time (different memory modules only). This facilitates high speed data transfer. Memory mapping, which requires the VORTEX II OS, provides extensions of main memory up to 256K in a page format. Memory protection is provided on a page basis. The translation procedure involved in memory mapping increases cycle time by about 150 nano-seconds. However, an interleaving procedure can be used to reduce cycle time by twenty-five to forty-five per-cent depending upon type of memory.

The configuration below is the minimum for support of the TOTAL Data Base System. Again prices are approximate. As mentioned above, TOTAL requires the VORTEX

II OS. This OS and the TOTAL software is priced at approximately $1,000 per copy.

Device	Pur price	Mo. maint.
CPU -V72-1100		
64K words	27,000	295
Moving-head disk - 2.34 M words	14,000	110
Tape unit and control	9,000	75
Medium speed line printer	10,000	130
CRT terminals (4)	12,000	20
Total	74,000	635

Detailed information on the system is contained in Appendix D.

TOTAL is based on a network data base organization. It utilizes direct linkages or threads to relate groups of data. TOTAL utilizes two types of files called Master and Variable files. Records within a master file may be independent or may be linked to records within a variable file. Variable records, on the other hand, must be linked to a master file either directly or indirectly through another variable record. One variable record may be associated with multiple master records. All linkages are stored within individual records. Each master record has a link to the first variable record in the file and to the last variable record in the file, ie., if a master has three variable files associated with it, six pointers will be present. Each variable record has pointers to adjacent variable records associated with the same master record.

TOTAL incorporates both a Data Base Definition Language (DBDL) and Data Management Language (DML). Both are English like using key phrase form. DML functions are limited to the opening and closing of files, serial

processing of files, program sign on/off, and data record logging. The inquiry and update functions must be written in a host language. Languages presently supported include FORTRAN IV, RPG II, BASIC and DASMR (Varian Macro Assembler). DML commands are issued via CALL statements in the selected host language. The DBDL is intended to provide data base independence with relationship to applications programs.

TOTAL is a reentrant foreground task. Hence multiple users can use the one resident copy simultaneously. TOTAL permits many users to access a data base in read-only mode with only one user allowed access in a read-write mode. Multiple users can addess different data bases in a read-write mode simultaneously.

3. Microdata REALITY

The last system to be described is the recently introduced Microdata REALITY. REALITY is a generalized data base management system based on a Microdata 1600 minicomputer, associated peripherals, and extensive software systems. REALITY is presently available through authorized dealers on a regional basis. The most significant machine feature of the system is the extensive use of micro-programming to implement both operating system and data base management functions.

Specific functions accomplished via the micro-code(firmware) are:

1. Virtual memory operating system,
2. Software level architecture,
3. Terminal input/output routines.

Virtual memory OS in firmware reduces the system monitor core requirement to 4K. Further, the resulting

demand page environment frees the user from core restraints and provides a program area in accordance with the capacity of the disk storage system.

The software level architecture includes facilities specifically designed and optimized for information management. Assembly language instructions implemented in firmware include character moves, searches, and compares of variable length fields and records. Also included is a set of programs for information management.

The input/output routines implemented in microcode enable a large number of on-line interactive terminals to interface with the CPU without degradation of response time. Block level control of communications between devices and buffering of messages at the block level permit the delay of software interrupt until a complete block transfer is finished.

The configuration below reflects that needed to support four users operating in a multiprograming mode on common or different data files. Included in the bundled system price are all firmware functions, several procedural languages and one inquiry language, a report formatting function, and a file/data security system.

Device	Pur price	Mo. maint.
CPU - 1600 Processor		
24K Bytes core	NA	NA
10 M Bytes disk	NA	NA
165-cps Printer	NA	NA
$10^1/^2$ in reel tape unit	NA	NA
4 data display terminals	NA	NA
Total	64,300	460

Since REALITY is a bundled system, price cannot be presented by component.

REALITY file structure is based on a set of dictionaries. The dictionaries have a hierarchial structure with the Systems Dictionary at the highest level. The Systems Dictionary contains legal log on names, passwords and security codes. Each entry in the dictionary contains a pointer to its corresponding lower dictionary, the User Master Dictionary. This dictionary contains user authorized vocabulary words, accessible file names, authorized procedure names, and a description of the structure of information within the dictionary. The next level dictionary is the User-File Dictionary. It contains the definition of the structure of the data in the user's file. The definition includes attribute names, sizes, access and display methods. Finally, there are pointers to the last and lowest level User-File Data Dictionary. This file contains the actual user data in variable length records. The fields within any record are also variable length allowing the storage of multiple attribute values per attribute name. Record size has an upper limit of 32K bytes.

Physical file structure is based on the virtual memory system. Files are stored randomly in 512-byte pages or frames. A hashing algorithm is utilized for addressing. Automatic provisions handle both the hashing collisions and frame overflows (record overlaps a frame boundary).

Data Base definition and creation is accomplished through the use of the systems on-line editor. The inquiry function is achieved via a generalized non-procedural inquiry language called ENGLISH. The update function is accomplished through the use of several procedural languages. An extended version of BASIC, called DATA/BASIC,

is included with the system. KEYSOURCE is a language developed and distributed by The Computer Works, the Northern California REALITY dealer. Both languages when used in conjunction with the system procedural facilities (PROC), provide a means for developing self-contained data base update facilities suitable for use by a non-programmer user.

ENGLISH, as mentioned previously, fulfills the inquiry function on the REALITY system. It utilizes fixed format sentences with the following syntactical form:

-VERB--FILE--ATTRIBUTES--SELECTION CRITERIA--MISCELLANEOUS CONNECTIVES.

Verbs are provided to list, sort, count, and select files in accordance with attribute and selection specifications. File and atributes are nouns used to specify desired fields and records. Selection Criteria specify the logical relations governing the retrieval while Connectives provide a means of combining and modifying the action of a verb. ENGLISH uses the dictionaries to authenicate user's privileges and to construct interfaces with the structure of specified files.

V. NETWORK CONSIDERATIONS

A. INTRODUCTION

The operational forces of the U.S. Marine Corps, the Fleet Marine Force (FMF), are supported in the data processing function by Force Automated Services Centers (FASC's). Each Wing and Division has an IBM S/360 computer located geopgraphically near the respective headquarters. The installations are tasked with supporting these organizations, mainly through the processing of Class I (Marine Corps wide) systems. Most of these systems are designed as information systems for the strategic manager and require extensive input from the operational level. Although the operational manager may request specific data processing support from the FASC, rarely does he get what he desires because of the priorities involved. The local manager is responsible for source data entry but does not reap the benefits of managerial assistance from the information system.

Although usually in a garrisonned situation, all operational units in the Marine Corps, usually down to the squadron/battalion level, must be ready to deploy expeditiously, perhaps for an extended period of time. Currently, data processing support ceases upon deployment and units must resort to a manual system. This reversion is not applicable to a Marine Amphibious Force (MAF) which is approximately a Wing plus Division size unit. In this case, one or both of the FASC computers is supposedly capable of being relocated to the operational area. However, such a relocation has never been attempted and the transportability of a S/360 system is questionable. Besides taking an extended amount of time to relocate, resulting in interrupted support, the ruggedness of the machine is also

questionable. Some other means of support have been proposed, one of which is the minicomputer network suggested in this thesis.

B. STATEMENT OF PROBLEM

It has already been mentioned that the FASC's are used mainly to process Class I systems. Examples of these Class I systems are:

1. Supported Activities Supply System (SASSY)
2. Marine Corps Unified Material Management System (MUMMS)
3. Mechanized Embarkation Data System (MEDS)
4. Marine Corps Automated Readiness Evaluation System (MARES)
5. Manpower Management System (MMS)
6. Force Status Report (FORSTAT)
7. Navy Maintenance and Material Management System (3M)
8. Division Logistics Data Base (DIVLOG).

Various forms of source data entry exist at the lower levels. It seems that each Class I system has its own input method and no thought was given to integration of all source data procedures. Since most of the Class I systems were designed and formulated around second generation technology, a proliferation of punched card input still exists. To get the punched card, direct keypunching and optically read character sheets are used. The JUMPS/MMS system uses a font typewritten form for source data entry. All these systems require a great deal of data manipulation at the source data level.

Although the lower level (squadron/battalion) manager is

responsible for submitting the source data, his feedback for validation purposes is quite slow. At least 12-hour turnaround time is required before the user is notified of an error. He then must correct the error, usually by regenerating a source data form, and resubmit for the next processing run. The user is usually judged on his error rate, yet he has virtually no control over validation of his submissions and the system is not very responsive.

The concept of transportability and data processing support has already been mentioned. Additionally, the geographic separation of units should be discussed. Even in a garrison situation, operational units of an organization may be distributed over a large geographical area. A good example is the Third Marine Aircraft Wing. Units of 3D MAW are scattered from El Toro to Santa Ana, California, to Camp Pendleton, to Yuma, Arizona. All these units must be supported by the FASC located at El Toro. Presently, source data from units not stationed at El Toro are sent by mail, by courier, or by AUTODIN. Such means of communication tend to offset the effectiveness of computerized systems and further the response problem.

Several various size units may be deployed depending on what the situation warrants. The MAF, the Marine Amphibious Brigade (MAB), Marine Amphibious Unit (MAU), squadrons, battalions and even company size units may be required to deploy. It is assumed, however, that the smallest units requiring data processing support are squadrons/battalions.

C. NETWORK PLAN

Two types of processing must exist at the local level due to requirements specified by Headquarters Marine Corps (HQMC) for integration of any new systems with existing systems: processing of systems designed to be implemented at the local level, such as a training information system, and front end processing for higher level existing systems.

The local processing would enable the lower level to manage its operation more effectively and give some freedom to this manager to develop some of his own applications. Several areas exist where local processing is pertinent because of the size of the data bases involved (small) and because of specialized applications at this level, such as training. It is also imperative that the local manager have some control over the data processing function if he is to feel that he can utilize it freely.

Of primary concern to the Marine Corps is source data automation, or how to more effectively enter data required for existing systems. Much of the error correction and validation is done by the large scale systems. A minicomputer as a front end processor, along with a CRT, would definitely enhance the source data enty process and provide immediate error detection and correction capabilities.

Although usually in a garrisonned situation, units must be able to take some data processing support with them if they are deployed. Their recourse is to revert to a manual system which becomes increasingly difficult as the data processing function becomes more entrenched in the Marine Corps. Automated systems are not necessarily set up the same as the manual system and the longer the automated system is used the less familiar personnel are with any

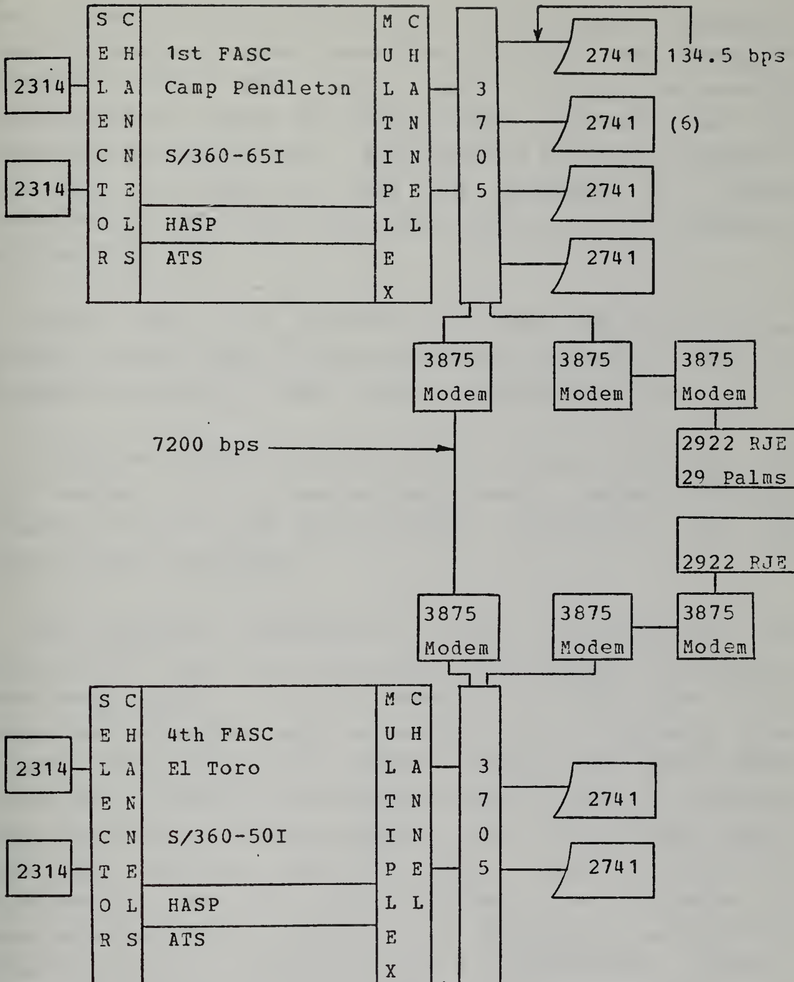
manual system.

Three physical restrictions must therefore be met by any proposed network. Data processing support must be available when garrisonned, when afloat or moving to a deployed area, and when actually deployed for any extended period of time. These restrictions imply certain capabilities that the network must have. if the computer is to be transported, it must be small enough and rugged enough to be moved easily. Some type of communications facilities are also required if source data is to be sent to the home site (FASC).

The network should be designed around the flow of information and not the organizational scheme. The organizational scheme is Division-Regiment-Battalion or Wing-Group-Squadron. The flow of information in most systems usually bypasses the intermediate level, or is consolidated but not processed at the intermediate level and forwarded for processing. As the network becomes more sophisticated as a management information system, it is conceivable that the highest level will expect certain reports from the lowest level, such as percentage of people not attending human relations training for a certain time period. These reports may be computerized to the extent that the request and subsequent processing can be accomplished with little or no human intervention and direct communications between computers. It is important that the intermediate commander be privileged to such information before the highest level commander so he is not surprised by any investigation of his units. An analysis of information flow must be made using extensive input from the intermediate level.

The current hardware configuration which supports FMF units is fixed for each MAF. Associated with each Marine Aircraft Wing (MAW) is a IBM S/360-50I and with each Marine

Division a IBM S/360-65I. These installations serve as the main processing sites and are positioned near the appropriate headquarters. A schematic of the configuration for I MAF is shown in Figure 13.



Current Hardware Configuration for I MAF

Figure13.

The internal memory requirements for the Houston Automatic Spooling Program (HASP) are 72K and for the Automated Terminal System (ATS), 56K. HASP acts as the control program for the remote job entry (RJE) terminals and the computer-to-computer link. ATS controls the 2741 terminals. Jobs may be routed to HASP for execution at either installation, but output from these jobs is at the computer, not the 2741.

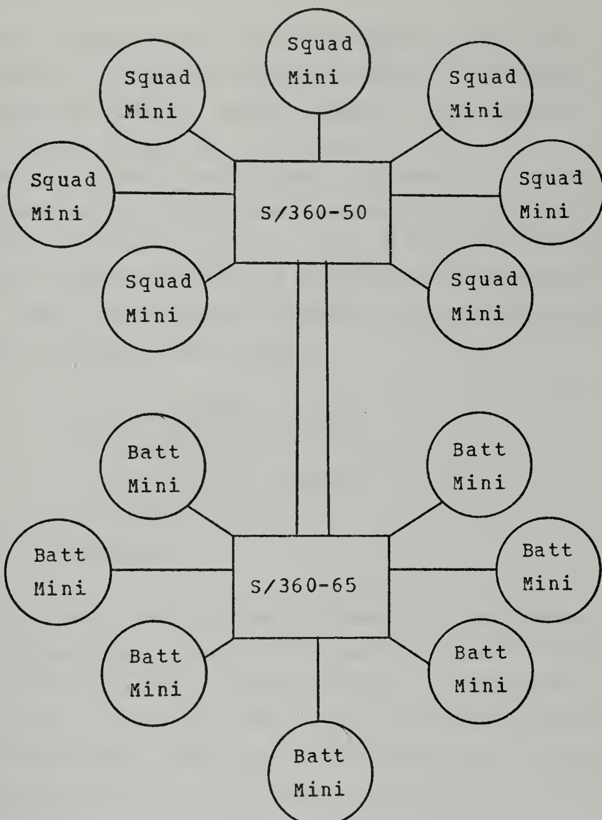
Source data may be entered on the 2741 and routed to the printer, punch, tape or disk at the host computer. This is independent of HASP. Note that all hardware is IBM.

A limited number of 2741 terminals exists and these are not located at any operational level. Rather, they are usually placed in some staff function associated with the higher level management.

The proposed minicomputer network configuration is designed to meet all software and hardware restrictions previously noted. It is based on the fact that squadrons/battalions are considered to be primary data sources for input into the present system and that these units are probably the smallest that will ever be deployed requiring data processing support. The intermediate level will be bypassed because this level's information requirements have not been analyzed. No sub-units of squadrons/battalions based in a different geographical location will be supported by a minicomputer system, although a CRT device may be resident at the sub-unit. This would mean additional communications lines from the sub-unit to the minicomputer.

The topology of the network will have a star configuration, since only the operational levels are required to communicate with the FASC's. Also, the local

units will probably be the smallest to deploy requiring data processing support. It is depicted in Figure 14.



Proposed Minicomputer Topology

Figure 14.

D. HARDWARE REQUIREMENTS

The hardware for the minicomputer system must meet both the physical and processing constraints discussed above. In the systems design phase, memory and peripheral equipment requirements will be developed for the stand-alone applications. Any specialized processing functions are also evaluated, such as data base management, file structure, and retrieval characteristics. Such a process was accomplished when the personnel training information system was designed. Based upon the training system, the following hardware was chosen to support this system:

* CPU expandable to 64K
* 2 Disk units
* 2 Tape units (backup)
* 4 CRT's
* Printer

Note that this hardware was chosen to facilitate only one application area. The 4 CRT's are for user convenience only and do not connote multiprogramming. Overlapped functions will exist allowing the CRT's to be used simultaneously. Further analysis must be undertaken to determine other applications which might possibly be done on the system.

Front end processing applications have not been evaluated. However, an example of front end processing is presented so that the applications may be considered.

The unit diary reporting system is the source data entry for the JUMPS/MMS. A clerk must use a font typewriter to type an OCR form. If he makes more than two typing mistakes in a row, he must restart the process. Forms must be lined

up perfectly in the typewriter. Once the form is typed, it is mailed or sent by courier to a computer for logical processing. If logic errors are detected, the user is notified and he must retype and resubmit the form. If the form is accepted, data is sent to Kansas City for furhter processing. Errors may again be noted and resubmission requested.

If a minicomputer network were implemented, the minicomputer could be used as a source data enty station. Using a CRT, the operator could enter information based on a pre-determined format presented to him on the CRT. If he notices a typing error, it is simple to correct it by backspacing. When he finishes one form, an edit is performed by the computer. If the form has logic errors, immediate correction can be made. If the form is accepted, it is spooled. When a batch of forms has been completed (daily), a send key is pushed and the data is transmitted to the larger machine for consolidation and future processing.

The above has been a simplified example of what a front end processor could do. Further analysis of applications must be accomplished.

E. COMMUNICATIONS

Computer-to-computer communications must be available for the network to function. Several restrictions will be put on the hardware because of this requirement and the general strategic plan for support to be implemented. Many minicomputers use a 16-bit word. Minicomputers frequently employ ASCII as the character code. The S/360, on the other hand, uses a 32-bit word and EBCDIC as a character code. If data is to be transmitted from computer to computer, conversion must be done at one site or the other.

Teleprocessing software exists for making the code conversions. The Marine Corps states that the maximum terminal data rate allowable is 2400 bps. Therefore, consideration should be given this restraint in relation to the following requirements.

While in a garrison situation, leased or hard wired communications are available. The cost of these lines and utilization times must be determined for each minicomputer site that is to be initiated. The system may be in use only 8 hours a day but a 24 hour a day capability may be required. This means that leased lines must be utilized.

While afloat, it is assumed that if the system is to be used at all, it will be used as a stand alone system. If afloat for an extended period of time, the Navy's Amphibious Support Information System (ASIS) can be used for data processing support.

When deployed for extended periods of time for which data processing support is needed (a policy decision), the need to communicate still exists. Hopefully, telephonic communication capabilities will be available. However, if they are not, then consideration must be given to radio/satellite links to accomplish the communication. Power requirements must also be met. The communications engineer should investigate these requirements and recommend solutions.

F. OPERATING SYSTEM

The operating system has previously been defined as a resource manager. For the network, it must also be responsible for communications. A priority interrupt

structure should be used at both computer sites. All I/O for communication should be batched and spooled so that efficiency in processing will be ensured. For instance, assume that the batch from the unit diary entries is ready to be sent. The 'send' function would activate an interrupt at the receiving computer which accepts the input and spools it for future processing. If this system is used, a half-duplex line is all that is needed for communication. However, a full-duplex line is recommended because of faster transmission rate, less noise, and relative cost (10 percent more than half-duplex).

G. IMPLEMENTATION CONSIDERATIONS

The intention of setting up an extensive array of computers is not to build a data processing empire. As simple a system of operation as possible is the goal. There will be no programming at the minicomputer site and no compiler available to the user. The operator will only need to have knowledge of how to turn on the system, mount tapes and disks, and be able to use the CRT. Although not as simple as using a typewriter, the use of the system should not require extensive training. The system is designed for use in the field by people unacquainted and untrained in the computer/data processing discipline.

It is envisioned that one or two small teams located with each MAF will exist. These teams will be staffed with the data processing expertise needed to implement and maintain the entire network of minicomputers. Any programming for additional needs will be done by this team. The team would also be responsible for training of operators so that schools could be held locally. No additional manpower need be used for these teams as they can probably be reassigned from the existing FASC Table of Organization.

Maintenance is a serious consideration for a network of this size. To reduce possible frustration, it is suggested that all hardware be purchased or leased from the same manufacturer. If the Marine Corps does not desire to train its own maintenance people, then a contract for maintenance should be made. Possibly, a maintenance man could be responsible for all minicomputer systems in a given geographical area, and only be concerned with Marine Corps systems. Upon deployment, the question of maintenance 2ecomes more complex and suggests a marine who is trained in this area.

The minicomputers and the network need not depend on IBM S/360 computers. It will be necessary to provide hardware and software compatibility between any new host computer and the minis as far as computer-to-computer communications is concerned.

H. COST

The cost of the network is difficult to measure because of the lack of knowledge of communications facilities needed. However, a projected cost can be made for the hardware itself. A single computer system would run approximately $50,000. If a number of computers are purchased, the price can be expected to decrease.

There are 169 squadron/battalion units in the FMF. Ground forces account for 72 battalions and air forces account for 57 flying and 40 non-flying squadrons. If hardware were purchased for all 169 units, a one-time cost of 8.45 million dollars would be incurred. No communications or maintenance costs are included in this figure, and many other costs should be evaluated.

Possible cost savings are difficult to analyze. According to the Informatics study, personnel savings accrued by automating a training system would be from 5.6 to 8.52 million dollars per year in the FMF. This savings would result from a redistribution of personnel. The release of personnel for other jobs would mean a reduction in recruitment to man those jobs. Another cost saving would result from paper, card and OCR form reduction.

Although the network would probably not save any money, the local manager's capabilities would be enhanced tremendously. The network would also provide a source data entry system for existing Class I systems.

VI. CONCEPT OF TRAINING INFORMATION NETWORK

A. DEFINITION OF A TRAINING INFORMATION NETWORK

1. Introduction

The previous sections have addressed the intra-unit training problem. A system has been described which provides for storage and retrieval of training information at the squadron/battalion unit level. The system inputs from outside the unit are FISMNPWR files. System outputs distributed outside the unit are the summary reports. The inter-unit exchange of this data has been viewed as a manual process. The preceding section in addition to training data considered other functions which are computerized. A network was then developed to replace the most frequently used paths of information flow with a digital communications link. This section explores a different approach by defining a network over which inter-unit training information can be exchanged throughout the organizational structure of the basic elements of the Fleet Marine Forces (FMF).

2. Terminology

The network to be developed parallels the organizational structure of the Marine Corps. Only the vertical flow of information is considered. No attempt is made to analyze the lateral flow of information between similar units at the same command level. This information exchange is both less rigidly structured and less frequently exchanged. The small amount of information being exchanged between like units does not appear to require digital exchange.

For ease of reference, the various organizations are described as follows: as levels as follows:

LEVEL 1	DIVISION/WING
LEVEL 2	REGIMENT/GROUP
LEVEL 3	BATTALION/SQUADRON

Thus the network structure to be defined can be thought of as the tree structures each of which is rooted at a level 1 unit. The level 1 unit is tied to several level 2 units each of which is in turn tied to several level 3 units. A further communications link (possibly manual) between level 1 units and a System/360 installation is assumed to exist. This link is not addressed in this thesis.

The following terms are used extensively in developing the concept of a network of systems:

1. Training Information System (TIS) - This is the squadron/battalion data base system developed in the previous sections.
2. System User - This is any individual within a unit who requests information from his own TIS.
3. Network User - This is any individual who requests or receives information from another unit's TIS.
4. List - This is one or more attribute/attribute value pairs. Lists specify retrieval parameters to be used in a search.
5. Reports - These are recurring training summaries required of subordinate level units.
6. Communications Link - This refers to the digital communications medium over which TIS's exchange information. The link is said to be established when two TIS's are on-line and capable of communicating.
7. Message - This refers to a formatted bit stream passed

over the communications link in order to request or provide information.

3. Benefits of a Network

At level 1 and level 2 there are personnel subject to the same training requirements as at level 3. Personnel assigned at level 1 or 2 must be the subjects of the same type of record keeping that occurs at level 3. There are additional training functions which are performed at these levels. They include the following:

1. Establish training requirements for subordinate levels.
2. Monitor progress of training through the periodic reporting system.
3. Report progress of training of all units at a subordinate level to a unit at a higher level.
4. Monitor the effectiveness of training programs of subordinate levels through inspections and training exercises.
5. Allocate training quotas among subordinate level units.

These functions can all be enhanced by the increased availability of data obtained by communicating digitally with the TIS's of subordinate units. The existence of a training information system in itself forces better organization of the various training requirements. This is particularly true if the organization requiring new training is required to implement the additional record keeping in its own system. A new file must be defined and possibly an old file deleted. The result will be that at some point the addition of new records must be balanced by the deletion of old ones. This is tantamount to forcing the deletion of old requirements when new ones are added. For example, if a new

training program is established at level 1 for all subordinate units, then the training officer at that level must not only be concerned with the promulgation of the requirement but also with the actual implementation of the record keeping in the training information system. The level 2 training officer must also consider this in passing the requirement to level 3. Clearly, each training officer, in managing the data bases, is managing the unit training program as well. The result is a type of forced review of all training requirements periodically, hopefully preventing a morass of outdated requirements with their attendant reports.

Closer monitoring of progress of training can be accomplished with a network. One reason for instituting training reports is to insure that subordinate units are in fact conducting the required training. The reports themselves may have no meaning beyond reflecting performance above some acceptable threshold. The preparer of the report may even schedule large amounts of training at the eleventh hour in order to be able to report favorable results before the reporting deadline. Once these results have been used to perpetuate the self-deception that an effective training program exists, they are filed away until required for an inspection. There is little alternative to this approach in a manual system. Periodic reporting is a sampling method employed by supervisors. If the samples were requested at irregular intervals, the result would be a truer measure of the ongoing effectiveness of the training program. This approach however is infeasible in a manual system. Further, it would likely be resisted by subordinates as oversupervision. If a network of training information systems existed however, the supervisor could obtain the information whenever he needed it. He could sample at random intervals without disrupting the work flow in subordinate units. He may optionally retain hard copies of

the information received from the subordinate units. The point is that he would be better able to evaluate the effectiveness and consistency of the training at the lower levels. Training records of the subordinate units would be open for inspection at all times.

Similarly, if level 2 were reporting to level 1, then the same random sampling technique could be used with the additional requirement that level 2 summarize information from all level 3 units before passing it to level 1. The reporting philosophy remains the same, the user of the information requests the data when he needs it.

The monitoring of the effectiveness of the training can be enhanced through networking the training information systems. A comprehensive training program established at level 2 can be closely monitored by periodically requesting from level 3 an array of information designed to diagnose weaknesses in the training program. Formal inspections of training records would be required less frequently. Inspections could then focus on observing performance.

Finally, there is another training function which could be facilitated by a network of TIS's. When school quotas become available at level 1 then all level 2 units can be polled to determine how many individuals have not attended the school. Level 2 units can then in turn poll level 3 units for the information. This assumes that each training information system defines files for school training information.

4. System Placement

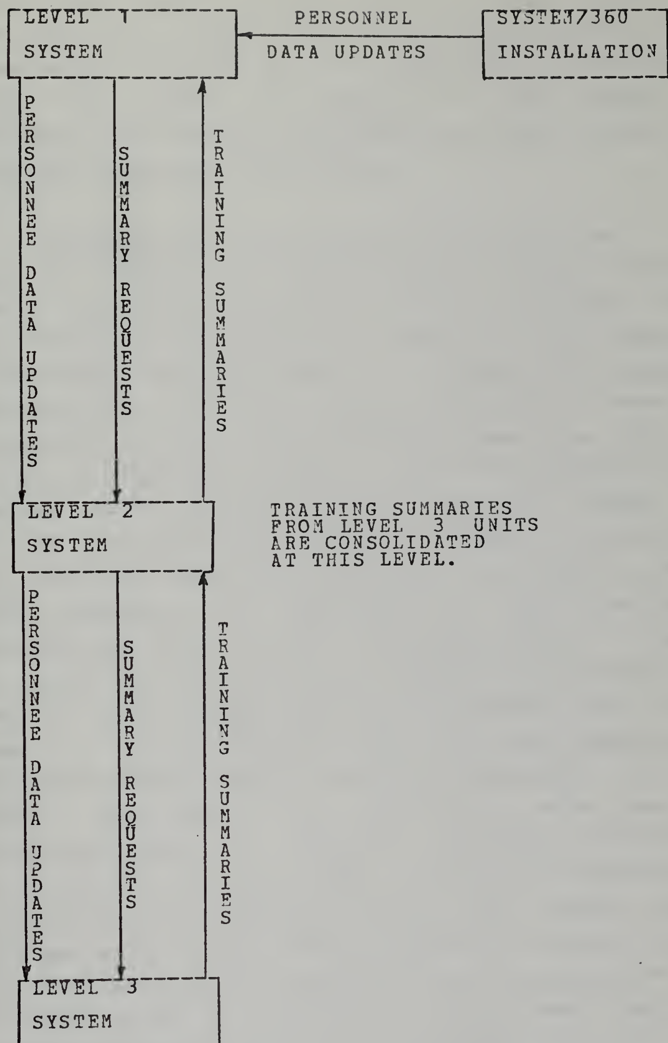
Because of the requirement to maintain basic training data on individuals throughout the Marine Corps, level 3 type TIS's must be present throughout the command

structure. It would be wasteful to dedicate a TIS solely to the consolidation of subordinate unit data. Therefore level 1 and level 2 systems may be used by both a unit (such as a Headquarters Unit) training officer and the training officer for that level. The systems should be designed so that any system is capable of handling both level 3 and level 1 and 2 functions. Physical location of units may then determine how many systems are required in garrison.

B. METHOD OF INFORMATION EXCHANGE

1. Categories of Information Required

The training information network provides for two types of information flow. It provides for the forwarding of training data from level 3 units to levels 1 and 2 and for the sending of verified personnel data. Verified personnel data is originated at a System/360 installation. It enters the network at level 1 and then is distributed to the appropriate unit. The following diagram depicts information flow in the network.



Network Information Flow

Figure 15.

Personnel data fields within master records cannot be updated directly by the user of the TIS at level 3. The data can only be changed at the System/360 installation. Whenever new files are generated at the System 360 installation, the data can be provided to level 1 units for distribution throughout the network.

Information flow from level 3 to levels 2 and 1 is summary in nature. The level 3 user may request from the system the names of all personnel who have not fired a rifle range during the calendar year. The level 2 user does not need the names, but may desire the number of personnel possessing that attribute value. The view is taken here that supplying names to the supervisory level can lead to the loss of authority at level 3. Thus, for example, it is considered a valid request if level 2 were to request the number of personnel who need to be sent to a particular school. This can lead to an allocation of school quotas. Specific assignment of individuals to fill those quotas remains with the level 3 unit. Conversely, level 1 and 2 users should have free access to summaries of training data at level 3. It is inappropriate for the level 2 user to be required to request from the commander of the level 3 unit permission to access summary training information. In the networked environment the subordinate units are not permitted the luxury of concealing training deficiencies until the unit has corrected them. Rather, the deficiencies may become known at subordinate levels only shortly before they become known to the training officer at the next higher level. Although such an arrangement may be annoying to the lower level unit, it nevertheless promises more hope for a coordinated, realistic and honest training program. The network alone, however, cannot solve problems unless the level 3 data bases are current.

2. Network Outputs

The network user requires the same types of output as the system user, reports of training accomplished and lists of numbers of individuals possessing certain attribute values. Although, as mentioned previously, the requirement for periodic reports should diminish as the use of random sampling increases. There will nevertheless be some recurring summaries desired by level 1 and level 2 users. These reports can be treated as pre-defined lists, simplifying the data transmission problem. Thus the user may request a basic report and receive data concerning PFT, Rifle Range, GMS Tests, Weight Control, etc.

The second category of requests, lists, is the more difficult capability to satisfy. It requires the network user to link with the general search functions of the TIS and then receive the summary results of the search over the network. The network user must therefore be able to request the information using the same retrieval language as the system user. Attributes and attribute values must be identified by exact field names. If the names of attributes are standard throughout the network, then the network user can be allerted at the time of input of an invalid field names.

The output from the periodic report requires only that the receiving system identify the type of report being sent by the responding unit, format the numbers which summarize the categories of information requested and output the report in a pre-defined format. Text for the names of attributes and values is retrieved from the pre-defined report description stored in auxiliary storage. The list output requires that the portions of the text to be output be included in the communications from the responding unit.

3. Network Requests

a. Frequency

The network user will provide fewer requests for information than the system user. The following table summarizes the estimated number of network communications weekly.

TYPE COMMUNICATION	LEVEL 1-2	LEVEL 2-3
PRE-DEFINED REPORTS	2	2
LISTS	10	5
PERSONNEL DATA UPDATES	5000	1000

In this table the communication "lists" for example is represented by 10 exchanges between level 2 and each of its subordinate units at level 3.

b. Response Times

Training information is not required on a real-time basis. Normal response time is 24 hours. Occasionally, "as soon as possible" requests are received in the manual system. These rapid response times are taken to be 15 minutes.

c. Request Delay

One additional problem occurs within the network. Because of the configuration and the random nature of requests, the level 2 system may not be able to respond to a level 1 request without first polling subordinate level 3 units. This means that there will be a delay in responding to some requests.

d. Amount of Information

Pre-defined reports can be transmitted as numerical data. The lists must be transmitted as a combination of numbers and characters. The following table provides an estimate (including redundancy) of the number of bits which must be transmitted daily. The average number of characters required to identify a list is taken to be thirty. Personnel data updates may be transmitted as a combination of characters and numbers. The transmission frequency is taken from the previous table.

	LEVELS 1-2	LEVELS 2-3
PRE-DEFINED REPORTS	50	50
LISTS	1500	750
PERSONNEL DATA UPDATES	1750000	350000
TOTAL	1751550	350800

The small amount of information to be transmitted and the lack of rigid time constraints on data retrieval indicate that a network for TIS's is not justified. The requests could easily be passed by telephone, the TIS queried, and the result returned by telephone. The cost of this manual handling would then be weighed against the cost of developing a network. Further, the dominating cost is that for transmission of personnel records, most of which do not change. If the System/360 only provided changed personnel data for updating purposes, the traffic could be reduced to 2 or 3 seconds per day. This thesis is, however, concerned with the development of a more general information system. The network concept therefore will be further developed in the belief that as the TIS is expanded to include other functions, the network load will increase to a point at which the benefits warrant the cost.

4. Batched On-line Inquiries

It is doubtful that the number of inter-system transactions for even a very general information system will ever be sufficiently large to justify the allocation of a dedicated communications link between units. The method of inquiry suggested is a batched on-line system which would opeate as follows:

1. When an inquiry is originated it is stored in auxiliary storage.

2. When a communications link is established all pending inquiries are transmitted.

3. When an inquiry is received a reply is prepared and transmitted to the requestor while the communications link remains established.

4. If the reply cannot be prepared within the time limit of the original communications link, it is stored in auxiliary storage for transmission during the next communications link.

The communications link is defined as telephone.

5. Communications Link

a. Messages

The information requests and responses constitute messages on the communications link. Each message should be independent. That is, there should be no processing which is suspended pending receipt of a reply message. Processing of a reply should not be contingent upon a request having been previously originated. Some of the network information, but not all, lends itself to fixed length formats. The exception is the request for lists. The following message types are defined to satisfy the

requirements of the network. The formats are provided solely to illustrate the message exchange concept.

Type 1 - Pre-defined Report

FIELDS

A. Message type - This specifies that the message is a type 1, Pre-defined report.

B. Report Number - This field identifies the number of the pre-defined report. Each Pre-defined report format used within a command is numbered by the using systems.

C. Request/Response - This specifies whether the report is a request or a response to a request. Special codes indicate that this is a message being returned to the sending unit because it is in error or failed parity checks. This field can also be used to indicate receipt of an error free message.

D. Originator- This specifies the unit originating the message.

E. Date Ending- This specifies the period ending for the report.

F. Data Fields - These fields contain the summary data for each attribute and value for which the report is defined. Numerical data fields are used.

Type 2 - Personnel Data Update

FIELDS

A. Message Type.

B. FISMNPWR Header Data - The data contained in the FISMNPWR Header is transmitted in a sequence of pre-defined fields.

C. Unit Name - This specifies the alpha-numeric designator of the unit to receive the information.

Type 3 -Lists

FIELDS

A. Message Type.

B. Request/Response

C. Attribute/Value/Number- A series of attributes and values are specified. When the message is a response, the number satisfying the attribute value is entered.

b. Communications Protocol

The communications link is a half duplex link established for brief periods of time to provide for rapid information exchange between two TIS's. The increased speed and reliability of full duplex communications may warrant its additional cost when the network traffic increases significantly. Some means of transmission discipline must be imposed. Each system should precede each message with a start message containing a unique code of consecutive 0's or 1's to denote the beginning of a communication. The start message is then followed by a fixed length Header message specifying the length of the Data Message to follow. Data messages are one of the types described above. Parity bits are computed for every eight data bits and inserted into the message. At the completion of all messages to be transmitted by one TIS the other TIS can then transmit. The operator must be able to designate that a system start communicating by either first transmitting or receiving. Additionally, the receiving system should provide an indication that it is ready to receive before transmissions of messages commence. The following sequence of events provides an example of message exchange under the communications link protocol.

1. Operator of level 2 system keys his system to transmit then receive. Operator of level 3 system keys his system to receive then transmit.
2. Level 2 sends a code denoting transmit phase commencing. This code can be part of a special Header Message.
3. Level 3 sends code for transmit acknowledge (ready to

receive).

4. Level 2 begins sending messages pairs consisting of a Header Message followed by a Data Message. The last header indicates that there no messages to follow.

5. Level 3 sends the code for transmit phase commencing.

6. Level 2 sends code for transmit acknowledge.

7. Level 3 begins sending message pairs. First, stored requests from the period of time preceding the establishment of the communications link are transmitted. Then responses to inquiries received from level 2 during the first part of the communications link are sent. The last header indicates that no messages follow.

C. SYSTEM MODIFICATION FOR NETWORK OPERATION

The system defined in a previous section would require the addition of a new function, the communications function, and the modification of several existing functions to incorporate a network capability.

1. Communications Function

The communications function can be keyed by the CRT operator or by the Operating System (OS). The operator can input either of the following:

1. Indication that a request message is to be assembled for transmission to another system.

2. Indication to begin network communications.

The former indication queues the message assembly processing; the latter, along with the indication to transmit or receive, is used to queue the OS to begin the transmit-receive cycle. Additionally, the operator must specify the unit to which messages are to be transmitted. The communications function is invoked by the OS when a message is to be processed. This is discussed in the

message reception processing.

a. Message Assembly

The assembly of messages requires that the sending system transform an operator request into formatted messages and store them for later output to the terminal equipment in accordance with the link protocol. The system must accept the input from the operator communications processing, build a message, construct a header message and store the messages in a queue awaiting transmission to the system designated by the operator. When the link protocol requires transmission the OS obtains the stored messages for output.

b. Message Reception

The system receiving a message must perform one of the following tasks:

1. Reply to an information request. - If the message is a request for information then a search must be conducted and a reply message assembled for return transmission. If the message type is pre-defined report, then the report number is used to index into a table containing the parameters to guide the search. If the message type is Lists, then the system must read the search parameters from the message as if they were being input by the system user. The searching is conducted as a request by the system user. At the completion of the search the number of finds is totalled. The message reply is then assembled and queued for transmission. If multiple request messages are received during one transmission, they are queued and processed in order of receipt. A copy of each incoming request and outgoing response is printed.

2. Process received information. - The receiving system in

this case is the original requestor. Prior to requesting information the network user must have identified a file in which information in pre-defined reports is to be stored. When the message is received, the report number is used to index into a table which contains the base address of the summary file to be updated. The index for the record number within the file is obtained from the unit number within the message. The fields are updated sequentially from the message. Lists are not filed, since by definition they are non-recurring requests. Personnel data updates are used to update the Master File. Master records not updated on three successive update opportunities can be purged from the system. The pre-defined reports and messages are both printed when received.

3. Route messages. - The only message routing defined in this system ocurs when a level 1 unit forwards Personnel data updates through a level 2 unit to a level 3 unit. In this case the level 2 unit must examine the unit name field in the message and place it into the queue for re-transmission to the correct unit. This is accomplished by finding the 7 character unit designator in a table and using the table index as the internal system unit number, enabling placement of the message into the proper transmission queue.

c. Validity Checking

Messages are not processed unless parity checks are successful. Additionally, there is a likelihood of user error resulting from incorrect specification of attributes or attribute values. The same validity checks are applied to the message requests for lists as are applied to system user list requests. When an error is detected, the message is retransmitted to the sender with the error field set. The transmit-receive protocol will insure that the error

message is sent during the same communication link. Received error messages are printed.

2. Operating System

The OS must be modified to provide for input and output over the communications link. Specifically, it must detect start codes, interpret Header messages and transfer complete Data Messages to auxiliary storage. After all messages have been input it must pass to the communications function the indication that messages are to be processed. For output the OS must retrieve the messages from auxiliary storage and continue to output them until all messages have been sent. Thereafter, it must output the header message denoting end of output.

At the completion of output the OS must shift to the input mode. On receipt of the header message denoting end of transmission the OS then shifts to the transmit phase, sending only the end of output header message if there are no messages to be sent. As long as the communications link exists the networked systems can alternate between transmit and receive. The OS must check parity on receiving data and compute parity bits and insert them into messages being transmitted.

Detection of the start of messages may be accomplished by hardware. In this case the OS would instead be required to interpret external interrupts from the terminal.

3. Operator Communications Function

The operator communications function must be modified so that for the list retrieval request the operator can additionally specify that the search is to be conducted

at the information system of another networked unit. That is, he can address a retrieval request to another unit. Four new sub-functions must be added to provide for the following:

1. Definition of pre-defined reports.
2. Addresses of units.
3. Keying of network communications.
4. Definition of summary files.

4. Pre-defined Reports

In each pre-defined report message the contents of the data fields represent the number of individuals possessing a specific attribute value. In order to interpret the reports, network users wishing to exchange reports must define them identically. A table in auxiliary storage is used to define the reports. The table item is indexed by the report number. Consecutive words within an item contain the exact field names for each attribute value for which information is desired.

5. Names of Units

Each unit in the network is assigned an internal system index to permit cross-reference between a unit's alphanumeric designator and the TIS files and queues. At system initialization the network user must enter the alphanumeric designator of his own unit and all other units with which his system will communicate. The information is maintained in a table and is referenced when Personnel Data update messages are rerouted or when a pre-defined report is to be filed. The table must also be referenced when summary files are printed in order to identify the unit with which the summaries are associated.

6. Keying of Network Communications

When the operator is ready to begin inter-system communications he must specify the unit to which messages are to be sent. Operator action may not be required if the terminal equipment itself provides for keying the communications by sending interrupts to the computer.

7. Definition of Summary Files

Summary files are defined by the user in the same manner as all other files except that the number of records is equal to the number of units in the network at lower levels. Thus when summary information is received from a unit with an alphanumeric designator stored in the second location of the unit table, it is stored in the second record of the summary file.

8. Printing of Messages

A print routine must be added to print the content of messages. It must examine the messages to determine if characters can be printed directly from the message or if the pre-defined report format must be referenced to obtain the characters representing the field. Numerical data must be converted to characters.

9. System Level Unique requirements

All systems regardless of the levels on which they function should be as similar as possible. Substantial differences in system design multiply software maintenance problems. There is some processing unique to certain levels. Levels 1 and 2 need to be able to keep a summary record on the personnel whose records are being maintained in the system. This could be accomplished by sending a

pre-defined report request to one's own system, conducting the search, building the reply message and then processing it. Additionally, level 1 systems must be capable of reading a FISMNPWR tape and building from it the Personnel data update for transmission to the correct unit.

10. Network Costs

The addition of the network capability to an existing system requires a terminal plus the software modifications described. The software changes are not extensive and could be added to the basic system without significant redesign. It is estimated that about 2K of programs plus another 2K of files and tables would be required. Additional auxiliary storage is required for the new function. The amount of main memory should not be affected because the network communications programs need to be core resident only during actual information exchange and not simultaneously with other existing functions.

VII. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

1. In designing a system to satisfy the training requirement it was discovered that there are very few short cuts that can produce an inexpensive system for training only. It is diffucult to assign to the training application a specific set of data elements and retrieval requirements. The training problem is very general. In order to create an effective system, provision must be made for retrieving on a number of keys and for storage of a wide variety of information. In solving the training problem it is necessary to create a flexible, general data base system which can be adapted to requirements peculiar to specific units. It is necessary to design a system which is not sensitive to the type of data being stored. Using this approach the addition of new squadron/battalion functions becomes largely a question of adding auxiliary storage. The additional auxiliary storage is relatively inexpensive. Any system developed for the squadron/battalion should be a general data base system to facilitate the expansion of system functions.

2. The actual cost of the system is difficult to determine. The cost of micro and minicomputers is decreasing. In the next few years more of-the-shelf software for minicomputer resident data base systems will become available. Given present trends in technology and data base system development, it is estimated that at some time in the future benefits will balance costs. The concept of a squadron/battalion level data base system should therefore be developed and the requirements of such a system should be further defined.

3. It is not likely that the use of an information system will result in decreased personnel costs at least initially. The training records must still be maintained even if they are automated. A competent individual will be able to spend less time keeping records or retried information. But, to balance this there will probably be more information kept and more requests for retrieval. Further, a competent individual must be selected to manage the training program, since the system will be is not properly maintained. Someone within the organization must perform the function of a data manager, keeping track of the system file configuration.

It is possible that, as functions are expanded, a lessening of the administrative work load will eventually permit the elimination of some clerks. It is also possible that the faster processing will permit the work load to increase or that the same size work force will be used in the automated system but with less overtime.

4. There are potentially large costs involved in training users and maintaining software. The former would require courses and perhaps several advisory teams throughout the Marine Corps. The latter would require a group of programmer/analysts to process software trouble reports, implement and test changes, and maintain system documentation.

5. Much of the design simplicity of the system designed in this thesis rests on the assumption that multi-programming is not required. Greatly expanding the functions may require a multi-programmed system. It is therefore critical to identify the expected growth of the system prior to designing a prototype system.

B. RECOMMENDATIONS

Throughout the development of the training information system which was defined in this thesis the need for user feedback became increasing obvious. It is difficult to determine if the system developed would be beneficial at all. Many system functions may never be used while other desirable features may not have been considered. It is therefore recommended that a system be developed and evaluated at a Squadron or Battalion by a project team. The system should provide for only the most rudimentary functions. It should be developed as quickly and cheaply as possible making maximum use of off-the-shelf software. This evaluation should produce the following results:

1. Determination of the types of functions which should be computerized.
2. The refinement of these functions into a detailed specification.
3. Determination of the human problems attendant to the system use, including suitability of input method, ease of system use and tendencies to introduce errors.
4. A measure of the speed required for information retrieval.
5. An estimate of the growth potential of the system.
6. An estimate of the potential network traffic.

Given the current advances in technology and the activities of vendors in developing data base systems the penalty for prematurely specifying a system inaccurately appears much greater than the penalty for over studying the problem.

APPENDIX A: DISCUSSION OF MARINE CORPS ORDERS AND PERTINENT TO THE TRAINING FUNCTION

MCO 1510.2H
Individual Training of Enlisted Marines

This order provides information, policy and implementing instructions for the individual training of enlisted Marines. It defines and discusses the various types of training existing in the Marine Training Program and their relationships to individual training.

The order outlines the commander's responsibility to accomplish mission oriented training, career training, essential subjects training and evaluation, and related training. Mission oriented training, as mentioned previously, is particular to each unit and is consequently not considered for generalization into the automated system. Related training augments individual training and includes human relations, troop information, etc.. These areas are discussed in detail subsequently.

Career training involves both Military Occupational Specialty (MOS) training and leadership training. Each marine may possess three MOS's indicating proficiency in a particular occupational field. MCO P1200.1B, The Military Occupational Specialties Manual, specifies the duties and tasks associated with each enlisted MOS according to rank. The descriptions include up to thirty task definitions. Because of the myriad of MOS's that exist, it is almost impossible for the training officer to monitor MOS qualifications. Rather, this function should be relegated to the lowest level (platoon/section). Therefore the following information should be included in the individual training record:

Field	Field Value
Primary MOS	4 digit MOS
Secondary MOS	4 digit MOS
Tertiary MOS	4 digit MOS
Billet MOS	4 digit MOS

Since MOS's change on a very infrequent basis, transaction rate would be limited in large part to updating qualifications. This normally occurs when a marine is promoted at which time he must demonstrate a higher level of MOS proficiency. Relatively speaking, the transaction rate per marine per year is negligible.

Career training also encompasses leadership training. The order specifies in detail the tasks and performance level marines must possess according to rank. The evaluation process is an extended process requiring a minimum of five hours. The training system must maintain information on the stage of leadership evaluation. The following information should be included;

Field	Value
Leadership stage	Stage
Leadership date	Date of last stage

Transaction rate associated with leadership evaluation scheduling and reporting is 10 transactions per marine year per 250 transaction days per year or, 0.04 transactions per marine per day.

Since only noncommissioned officers will be evaluated this figure is higher than normally would be expected, but is good for worst case.

Essential Subjects Evaluation Traning is assumed to be conducted on an annual basis. If a marine demonstrates

proficiency, i.e. passes the evaluation test, he is exempted from further training in the particular subject. The training system must reflect the particular essential subjects and an associated qualification. The following should be included in the individual training record;

Field
Code of Conduct/UCMJ
History, Customs, and Courtesies
Close Order Drill
Interior Guard
First Aid/Field Sanitation
Uniform Clothing and Equipment
NBC Defense
Service Rifle
Service Pistol
Individual Tactical Measures

The field value for each of these fields is qualification digit.

Transaction rate assuming a fifty per cent failure rate is 30 transactions per marine year per 250 transaction days per year or, 0.120 transactions per marine per day.

The order also suggests an individual training record for a manual system. There are certain information items on it which are pertinent to the trainning system but which are not discussed elsewhere. These should be included on the training record. They are:

Field	Field Value
Gas Mask Size	Size Number
Swimming Qualification	Qualification Code

MCO 1510.25A
Marine Corps Troop Information Program

This order provides guidance to commanders on the formulation and conduct of the Troop Information Program within the Marine Corps.

The Troop Information Program encompasses topics which augment the training required by MCO 1510.2H, Individual Training of Enlisted Marines. this training supports the primary Marine Corps mission. The program includes as a minimum the below listed topics:

1. Drug Abuse
2. Alcohol Abuse/alcoholism
3. Equal Opportunity
4. Personal Affairs
5. UCMJ
6. Character and Moral Education
7. Citizenship
8. Personal Conduct
9. Human Relations.

The training aspects of the Drug Abuse and Human Relations Programs are elucidated in directives which are discussed subsequently. The remaining topics require additional discussion.

The training associated with the other topics requires presentation of the topic but does not involve any type of evaluation/testing. Hence, the information required would consist of associating type of training by topic with those marines requiring training. Such information would be used to develop schedules, ad hoc reports, and would be a part of an individuals training record.

The following data elements should be incorporated into the training record:

1. Alcohol Abuse and Alcoholism-date

2. Equal Opportunity and Citizenship-date

3. Personal Conduct/character and Moral Education-date

4. UCMJ-date

5. Personal affairs-date.

Some closely related topics have been combined into one field on the assumption that they could be combined into one period of instruction. The field value would be the date the instruction was given to the individual.

In order to develop a transaction rate it is assumed that each topic would be presented on an annual basis. Hence, the transaction rate is: 10 transactions per marine year per 250 transaction days per year or, 0.040 transactions per marine per day.

MCO 3574.2E

Marksmanship Training With Individual Small Arms

This order establishes Marine Corps policy concerning marksmanship training with individual small arms. Every marine must be qualified and trained in the individual weapons associated with his grade and duties.

Consequently marines must receive instruction on and qualify on a marksmanship range with individual weapons. The two activities normally occur during the same time period. The individual weapons include the service rifle, pistol, and shotgun. The pistol and rifle are normally fired for qualification whereas the shotgun is normally fired for familiariztion only.

The following data elements are to be included in the automated individual training record:

Rifle-Qualification Score and Date

Pistol-Qualification Score and Date
Shotgun-Date (date familiarization course fired).

Marines are required to qualify on an annual basis. The transaction rate for recording is estimated at 6 transactions per marine year per 250 transaction days per year or, 0.024 transactions per marine per day.

Rifle and pistol qualifications are also currently recorded in the marine's administrative record (OQR/SRB). This suggests that it might be proper to record past qualifications in an auxiliary historical record. Since this data would be accessed very infrequently if should not be a part of the main record.

MCO 5100.19A
Marine Corps Traffic Safety Program
For Off-duty Military Personnel

This order provides information and instructions for the conduct of a Traffic Safety Program for off duty military personnel. The order requires all enlisted personnel under 25 years of age to complete a recognized driver improvement course. Further, those marines in this category must be given the course within the first six months of duty at their first permanent activity or unit.

At most installations the course is conducted at a centralized location. Hence, the training system has to identify those who must take the course.

Therefore it is recommended that a field for driver improvement be included in the training record. The value of the field would be the date the course is completed (mmyy). Transaction rate is estimated roughly by estimating that one half of all marines qualify for the course. Hence,

transaction rate is 1 transaction per marine year per 250 transction days per year or, 0.004 transactions per marine per day.

MCO 5350.4A Human Relations Program

This order establishes a Human Relations Program for all marines. The program is designed to enhance the combat effectiveness of marines through more effective relationships among marines and between marines and individuals outside the Marine Corps.

The program is conducted according to 'year periods' with the following breakdown:

1. First year twenty hours of orientation and small guided group discussion.
2. Second year twenty hours of review of year one program and the development of an individual action program.
3. Third and subsequent years twenty hours devoted to review of concepts of previous years program plus definition of and action on local human relations problems.

Since marines must take the courses in sequential order, all three phases of training will be conducted during the same training period. The training will most likely be broken into a maximum of ten two hour periods. Hence, the system must be capable of identifying the phase the marine is in and the number of hours training he has received during the current annual period. The following information should be included in the individual training record:

1. Date completed first year program
2. Date completed second year program
3. Date completed third and subsequent year program
4. Hours completed in current program
5. Date of last training session.

In addition the order requires that each unit select marines for training as Unit Discussion Leaders. This suggests the inclusion of a field in the individual training record with descriptor Unit Discussion Leader and descriptor value of date designated.

The order also requires that commands submit a Command Human Relations Chronology Report on a semi-annual basis. The information necessary to generate this report is a subset of that listed above.

Transaction rate for human relations is based on the assumption of the scheduling and recording of ten two hour classes per year. hence, transaction rate is 20 transactions per year per marine or, 0.080 transactions per marine per day.

The order requires that entries be made in the administrative records of marines indicating the completion of annual training and designation as unit discussion leaders.

MCO 6100.3F
Physical Fitness and Weight Control

This order establishes the Marine Corps policy concerning physical fitness of marines. The implementation of the policy requires that all marines participate in an effective physical conditioning program on a regular basis.

The order requires a minimum of three hours per week to be devoted to physical fitness. In addition all marines 45 years of age and under are required to pass a physical fitness test on a semi-annual basis. Those marines who either fail the test or are judged to be overweight are placed on a weight control and remedial physical fitness

program.

Marines serving in a combat, combat support, or service support unit within the Fleet Marine Force are required to pass a unit endurance test. This test is also conducted on a semi-annual basis.

The following data elements should be incorporated into the individual training record:

1. Date PFT test taken
2. Number of pullups completed
3. Number of situps completed
4. Time of running three mile run
5. Date unit endurance test taken
6. Qualification achieved on unit endurance test
7. Date placed on weight control.

An estimate of the transaction rate for this portion follows. It assumes an average of six events per year per marine based on required test plus a failure rate. The transaction is 12 transactions per year per marine or, 0.048 transactions per marine per day.

It is also recommended that physical fitness test results be recorded in a historical record.

MCO 6710.1B
Marine Corps Drug Abuse Control Program

This order establishes a program within the Marine Corps for the prevention and control of drug abuse in accordance with Department of Defense guidelines.

The bulk of the order is concerned with administrative functions outside of the realm of the training system. However the order does direct that enlisted marines receive

initial drug abuse instruction and overseas drug abuse instruction. Hence, the training system must identify those marines who require training. The following information should be included in the individual training record:

1. Date completed initial drug abuse instruction
2. Date completed overseas drug abuse instruction.

Since training is on a one time basis, the transaction rate per marine is minimal. However, assuming that this training is conducted on a monthly basis, scheduling events occur on a monthly basis.

The order requires that a record be made of this instruction in the marine's administrative record.

Marine Corps Institute

The Marine Corps Institute (MCI) correspondence training program is another area in which computerizing records could be beneficial. There is a need to closely monitor student progress. MCI furnishes a Unit Activity Report monthly to each unit showing the progress of students currently enrolled in courses. Upon receipt of this list the unit's training officer contacts those delinquent in course submission. Ideally, the student should be encouraged to submit lessons while he is still interested in the subject. After he has put the course aside for a month, it is doubtful that he will attack it with enthusiasm when confronted with a delinquency report. A more aggressive program to monitor student progress could be aimed at preventing the student from becoming delinquent. This requires extensive record keeping to ensure that periodic submissions are taking place. By monitoring student progress on a computer, lists of students who fail to meet the submission deadlines can be generated periodically, thus pushing the student into a rhythmic submission pattern.

Computerizing the system would not only provide a savings in record keeping but may well result in better learning. The information required would be course number, last submission date and an indication that the student has completed lessons and is waiting to take the examination or a re-examination. A separate list of all completions could be maintained.

However, the MCI courses were not considered for computerization in this thesis.

APPENDIX D - SAMPLE OUTPUT

MOS AND LOCATION REPORT
3RD MARINE AIR WING

PAGE [illegible]
AS OF 052073

LAST NAME	INITS	RK	SSN	PRI MOS	1ST ADD MOS	2ND ADD MOS	BMOS	EAS	DCTB	SCTY CLNC	PEBD	DOR	DAUS DEPN RESTR	COMP GCT	R	STR CAT	# OF DEPN	MCC	RUC	UD NO	EOD

MOS	08	07	06	05	04	03	02-1	WO	TOTAL
6702	0	0	0	2	7	2	0	0	11

LAST NAME	INITS	RK	SSN	PRI MOS	1ST ADD MOS	2ND ADD MOS	BMOS	EAS	DCTB	SCTY CLNC	PEBD	DOR	DAUS DEPN RESTR	COMP GCT	R	STR CAT	# OF DEPN	MCC	RUC	UD NO	EOD
[illegible]	[illegible]	[illegible]	[illegible]	6704	0430	0000	9910	UAO	721028	17	511001	620701	721019	139	C	4	04	143	[illegible]	[illegible]	[illegible]
CLARK	[illegible]	33	[illegible]	6704	0000	0000	6704	000000	720706	27	651116	710301	720612	137	C	1	01	144	22961	111	[illegible]
[illegible]	[illegible]	33	[illegible]	6704	2502	0000	6704	000000	720525	19	630923	700101	700909	000	C	0	02	144	22961	[illegible]	[illegible]
[illegible]	JR	32	[illegible]	6704	0102	0130	6704	000000	710903	19	590525	681101	690116	132	C	0	05	144	[illegible]	CCM	[illegible]
[illegible]	[illegible]	33	[illegible]	6704	0130	5702	6704	000000	720506	27	530623	660120	720401	128	C	0	06	144	22951	CCM	[illegible]
[illegible]	JT	33	[illegible]	6704	6502	6002	6602	000000	730327	17	631021	660803	730310	150	C	0	03	145	31020	[illegible]	[illegible]
[illegible]	JJ	33	[illegible]	6704	6710	0000	6704	000000	700724	19	650507	710101	[illegible]	132	C	0	00	186	[illegible]	CCM	[illegible]
[illegible]	[illegible]	33	[illegible]	6704	6733	0000	6402	000000	710715	27	600601	720601	710605	137	C	0	04	143	[illegible]	CCM	[illegible]
ARMSTRONG	RJ	32	[illegible]	6704	0000	0000	6704	730715	710308	27	690410	710903	000000	119	C	0	02	144	22961	CCM	[illegible]
[illegible]	FL	32	[illegible]	6704	0430	0000	6704	000000	710723	27	690321	710903	000000	123	C	4	01	144	22961	[illegible]	[illegible]
[illegible]	AJ	32	[illegible]	6704	0000	0000	6704	730901	710436	27	670223	710903	000000	120	C	0	01	144	22961	[illegible]	[illegible]
[illegible] JR	[illegible]	32	[illegible]	6704	0000	0000	9910	720701	710212	27	661216	710401	000000	135	C	0	02	143	31079	CCM	[illegible]
[illegible]	[illegible]	32	[illegible]	6704	0000	0000	6704	730901	710505	27	681031	710903	000000	132	C	0	02	144	[illegible]	[illegible]	[illegible]
[illegible] JR	[illegible]	32	[illegible]	6704	0000	0000	6704	731215	710810	27	690124	720201	000000	131	C	0	00	144	22961	CCM	[illegible]
[illegible]	[illegible]	32	[illegible]	6704	6710	0000	9412	000000	730511	19	681030	700701	000000	135	C	0	00	143	35160	CCM	[illegible]
[illegible]	[illegible]	32	[illegible]	6704	[illegible]	0000	6704	730715	710511	27	681202	710401	000000	122	C	0	00	144	22951	CCM	[illegible]
[illegible]	[illegible]	32	[illegible]	6704	0000	0000	6704	[illegible]	720323		660919	721209	000000	000	C	4	00	144	22464	CCM	[illegible]
[illegible]	[illegible]	32	[illegible]	6704	0000	0000	6704	731101	701218		680320	710701	000000	156	C	0	02	144	22961	CCM	[illegible]
MILLER	[illegible]	31	[illegible]	6704	0000	0000	6704	741101	720627	27	700617	711101	000000	000	C	0	00	144	22964	CCM	[illegible]

MOS	08	07	06	05	04	03	02-1	WO	TOTAL
6704	0	0	0	0	1	7	11	0	19

APPENDIX C: MASTER FILE RECORD DESCRIPTION

FIELD NO.	FIELD ID	FIELD TYPE	FIELD LENGTH
1	Social Security Number (SSAN)		
	First Digit (Service)	A/N	1 chars
	Remainder	N	9
2	Date of Rank (DOR) (yymmdd)	N	6
3	Primary Military Occupational Specialty (PMOS)	N	4
4	First Additional MOS (1MOS)	N	4
5	Second Additional MOS (2MOS)	N	4
6	Billet MOS (BMOS)	N	4
7	Expiration of Active Service (EAS)	A/N	6
8	General Classification Test Score (GCT)	N	3
9	Date Current Tour Began (DCTB)	N	6
10	Security Clearance	A/N	2
11	Pay Entry Base Date (PEBD)	N	6
12	Date Arrived U.S. Dependents Restricted (DAUSR)	N	6
14	Number of Dependents (NDEP)	N	2
14	Reporting Unit Code (RUC)	A/N	5
15	Unit Diary Number (UDNO)	A/N	3
16	Rotation Tour Date (RTD)	N	6
17	Pace	A/N	1
18	Strength Category (STR CAT)	A/N	1
19	Estimated Date of Departure (EDD)	N	6
20	Date of Birth (DOB)	N	6
21	Duty Status Code (DSC)	A/N	1
22	Quota Serial Number (QSN)	A/N	6
23	Orders Flag	A/N	2

FIELD NO.	FIELD ID	FIELD TYPE	FIELD LENGTH
24	Civilian Education	A/N	5
25	Service Schools (Total of 8)		
	Code-3		
	Year Attended-2	A/N	40
	Length of Header		145 **
26	Swimming Qualification	A/N	2
27	Gas Mask Size (GMS)	A	1
28	Traffic Safety Program (TSP)		
	(mmyy)	N	4
29	Marksmanship Training (MRKT)		
	Rifle (mmyy,qual(3))	N	7
	Pistol (mmyy,qual(3))	N	7
	Shotgun (mmyy)	N	4
30	Human Relations Program (HRP)		
	First Program (mmyy)	N	4
	Second Program (mmyy)	N	4
	Third Program (mmyy)	N	4
	Current Program Hours		
	Received	N	2
	Unit Discussion Leader		
	Date Designated (mmyy)	N	4
31	Physical Fitness and Weight		
	Control		
	Date Test Taken (mmyy)	N	4
	Pullups Score	N	2
	Situps Score	N	3
	Run (minminsecsec)	N	4
	Unit Endurance Test Date		
	(mmyy)	N	4
	Date Placed on Weight		
	Control Program (mmyy)	N	4
	Length of Trailer		64 **
	Length of Record		209 ***

APPENDIX D: MACHINE SPECIFICATIONS

Digital Equipment Corporation PDP - 11 Series.

1. DATA FORMATS.

The PDP-11 uses a 16 bit word. Operands may either be of single byte length or of one word. An option is available for floating point arithmetic which uses a 32 bit word. There are 75 instructions ranging in size from one to three words. Addressing is limited to 64KB without memory management option. With memory management, addressing is extended to 124K words. Eight addressing modes involving combinations of indexed, indirect and increment methods are available. Internal code is stored in ASCII format.

2. MAIN STORAGE.

Main storage is core with optional MOS or bipolar memory. Core cycle time is 950 nano-seconds. Parity checking is optional. Protection is provided through the optional Memory Management module.

3. CENTRAL PROCESSOR.

Eight user accessible 16 bit registers are provided. Six of the registers are general purpose with one for a program counter and another as a stack pointer. Indirect addressing is standard.

Instruction breakdown is as follows:

16 arithmetic
21 branch
 7 trap and interrupt

19 data manipulation

7 logic instructions

Instruction timing for the PDP 11/40 follows:

load/store	2.42/2.24
add/subtract	2.66/2.80
multiply/divide	9.66/11.30
compare/branch	2.75

Interrupts are based on a four level automatic priority system. Each level can be used for multiple, different priority peripheral devices.

The memory management option allows access of 128K words of main memory in a virtual storage system based on 32K word segments. Memory protection is provided through the use of bounds registers and status registers. Memory management allows two modes of operation. In the kernel mode, memory mapping and protection may be modified. In the user mode each user has access only to his memory area.

An instruction stack capability limited only by the size of main memory is standard. This eases the execution of reentrant routines.

4. INPUT/OUTPUT CONTROL.

The UNIBUS is the common device for all data access and transfer. All peripherals are treated in the same matter. Priority is established according to the location of the device on the UNIBUS. There is no logical limit to the number of devices attached to the bus. Access to the bus is controlled through the interrupt system. Maximum data rate over the UNIBUS is 2.5M words/second. All data transfers operate in a master slave mode.

5. PERIPHERALS.

Digital Equipment Corporation offers a wide variety of peripherals for its systems. The peripherals listed below correspond to the devices specified in the equipment list of the text.

RK05 DECDISK fixed head disk drive

- 1.2 M words per drive
- 70 milliseconds access time
- 90.25 K words/second transfer rate
- maximum of eight drives per controller

TM11 Magnetic tape transport and control unit

- 45 ips tape speed
- 9 track - 800 bpi density
- 7 track - 200/556/800 bpi tape density
- 36 KBS data transfer rate

LP11-J Medium speed printer

- 132 positions
- 64 characters
- 300 lines per minute

VT05B - A/N CRT display terminal

- 72 character
- 20 lines
- 110 to 2400 bits per second transfer rate

Varian Data Machine V 70 Series.

1. DATA FORMATS.

The Varian 70 series utilizes a 16 bit word with operands of byte or word size. A floating point option uses a 32 bit word. Instructions are either one or two words long. Addressing modes include direct, relative, indexed, or indirect. Internal code is ASCII.

2. MAIN STORAGE.

Main storage is core with optional semiconductor memory avaiable. Another memory option allows dual ported modules. Writable Control Store is another option available in bipolar form with a 190 nanosecond sycle time. Cycle time for core is 660 nanoseconds while that of MOS is 330 nanoseconds. Utilization of a memory map module allows up to 256 words of main memory. A limited form of wrap around strorage protection is provided via the memory map module. Writable Contol Store is attached in 512 word modules with a maximum of 2 K words.

3. CENTRAL PROCESSOR.

There are 16 general purpose user accessible registers. Other registers include a operand register, program counter, shift counter, processor key register, I/O key register, memory address register and I/o register. Indirect addressing is allowed to multiple levels.

Instruction breakdown is as follows:

18 load/store

14 arithmetic

10 logic
12 shift/rotate
30 register transfer/modification
21 jumps
18 jump and mark
18 execution
4 control
14 I/O

This is a total of 159 basic instructions.

Instruction timing is as follows:

	CORE	MOS
load/store	1.32	0.66
add/subtract	1.32	0.66
multiply	5.20	4.70
divide	5.57	5.07
compare,branch	2.06	1.03

A priority interrupt module (PIM) allows eight levels of priority interrupts which can be expanded to sixty-four levels. Each level possesses a unique memory address.

4. INPUT/OUTPUT CONTROL.

Three types of I/O operations are permitted over the party line, time shared bus. They are:

1. Direct memory access (DMA). DMA utilizes a cycle stealing sequence to transfer blocks of data at rates up to 330 K words/second. Higher speed DMA devices are available which increase the rate to 1 M words/second.
2. Priority memory access (PMA). PMA bypasses the I/O bus and processor allowing data transfers up to 1.5 M words per second.
3. Program controlled I/O. In this mode seperate program instructions are used for each character or word

transfer.

5. PERIPHERALS.

The peripherals listed below correspond to the devices specified in the equipment list of the text. Varian Data Machines also offers a wide variety of peripherals which are not listed below.

620-36 disk memory and controller
- 2.35 M words on one removable and one fixed disk pack
- 35 millisecond average access time
- 92 K words/second transfer rate
- maximum of two drives per controller

620-32 Magnetic tape transport and control unit
- 37.5 ips tape speed
- 9 track - 800bpi density
- 30 KBS data transfer rate

7067-21 Medium speed printer
- 136 positions
- 64 characters
- 300 lines per minute

E-2250D A/N CRT display terminal
- 64 character
- 20 lines
- 4800 bits per second transfer rate

Microdata Corporation REALITY System.

1. DATA FORMATS.

The Microdata 1600 series utilizes an eight bit word with operands of one to four bytes. Instructions are 16 bits. Internal code is ASCII.

2. MAIN STORAGE.

Main memory is core with 1 usec cycle time. Control memory exists in one of three forms. BROM is bipolar read only memory. PROM is programmable read only memory while ACM is alterable contol memory. All three possess a 200 nanosecond cycle time. Parity checking can be accomplished only through a software routine. No storage protection is incorporated into the hardware. Core memory is limited to 65 K bytes while ROM is limited to 1,792 words.

3. CENTRAL PROCESSOR.

There are two sets of 15 general purpose 8 bit registers. In the first set 6 are user accessible while all fifteen are accessible in the other set. Indirect addressing is available to one level.

Instruction breakdown is as follows:

- 17 conditional jumps
- 16 control
- 12 arithmetic and logical shifts
- 19 register to register
- 20 memory reference
- 8 stack control
- 6 input/output

5 character/string manipulation
2 decimal arithmetic

This is a total of 105 instructions.

Instruction timing is as follows:

load/store	4.6/5.2
add/subtract	4.6/4.6
multiply/divide	56/72
compare/branch	5.6/4.0

Interrupts exist in three forms, internal priority, I/O peripheral device, and individual external interrupts. Up to 128 interrupts are available. Each interrupt has a unique memory address and priority assignment.

4. PERIPHERALS.

The peripherals listed below correspond to the devices specified in the equipment list of the text.

2856 disk memory and controller
- 10 M bytes on one removable and one fixed disk pack
- 75 millisecond average access time
- 200 K bytes/second transfer rate

2815 Magnetic tape transport and control unit
- 25 ips tape speed
- 9 track - 800 bpi density
- 20 KBS data transfer rate

medium speed printer
- 132 positions
- 64 characters
- 130 lines per minute

CRT display terminal
- 64 character
- 24 lines

up to 9600 bits per second transfer rate

BIBLIOGRAPHY

1. Auerbach Publisher Inc., <u>Auerbach Guide To Minicomputers</u>, April, 1975.

2. Datapro Research Corporation, <u>Datapro Reports On Minicomputers</u>, 1975.

3. Davis, G., <u>Management Information Systems</u>, McGraw-Hill, 1974.

4. Informatics TR-74-1582-D6 Technical Report, <u>Data Management Device Requirements Study</u>, by Informatics Inc., 30 June 1974.

5. Katzan, D., <u>A PL/I Approach To Programming Languages</u>, Auerbach, 1972.

6. Modern Data, <u>1975 Minicomputer Market Survey</u>, February, 1975.

7. Rand Corporation Report R-1279, <u>Minicomputer: A Review Of Current Technology, Systems, And Applications</u>, by Dennis Hollingworth, July 1973.

8. Sammet, J.E., <u>Programming Languages: History And Fundamentals</u>, Prentice-Hall, 1969.

9. Smith, G.W., <u>ADS Support For The FMF, Briefing Presented To Headquarters Marine Corps Chief Of Staff Committee</u>, December 1974.

10. Systems Data Corporation Technical Memorandum - 2314000000, Planning Guide For Computer Program Development, by L. Farr, V. Labolle and N. Willmorth, 10 May 1965.

11. Systems Data Corporation Technical Memorandum - 322500001, Management Handbook For The Estimation Of Computer Programming Costs, by E. Nelson, 20 March 1967.

12. United States Marine Corps Automated Data Systems Plan, Fiscal Years 1975-1980, 1974.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Department Chairman, Code 72 Department of Computer Science Naval Postgraduate School Monterey, California 93940	1
4.	Professor N.F. Scheidewind, Code 55Ss (advisor) Department of Operations Research/Administrative Science Naval Postgraduate School Monterey, California 93940	1
5.	LtCol R.J. Johnson, USMC Head, Systems Planning Review and Implementation Branch Information Systems Support and Management Division Headquarters, Marine Corps Washington, D.C. 20370	1
6.	LtCol R.E. Jamison, USMC Marine Corps Liaison Representative, code 55Jm Naval Postgraduate School Monterey, California 93940	1

7. Major Robert A. VanHouten, USMC (student) 1
3207 Tallmon St.
Marina, California 93933

8. Capt R.P. Hansen, USMC (student) 1
Marine Detachment
Naval Missile Center
Point Mugu, California

9. Capt Robert E. Sonnenberg, USMC (student) 1
1260 Clearmont
Boulder, Colorado 80303

Thesis 161127
V169 Van Houten
c.1 A design of a personnel training information system for the U. S. Marine Corps.